(19) Europäisches Patentamt
European Patent Office
Office européen des brevets

(11) Publication number: 0 251 446
A2

(12) EUROPEAN PATENT APPLICATION

(21) Application number: 87303761.8

(22) Date of filing: 28.04.87

(51) Int.Cl.⁴: C 12 N 15/00
C 12 N 9/54, C 12 N 1/00

(30) Priority: 30.04.86 US 858594
06.04.87 US 35852

(43) Date of publication of application:
07.01.88 Bulletin 88/1

(84) Designated Contracting States:
AT BE CH DE ES FR GB GR IT LI LU NL SE

(71) Applicant: GENENTECH, INC.
460 Point San Bruno Boulevard
South San Francisco California 94080(US)

(72) Inventor: Wells, James Allen
64 Otay Avenue
San Mateo CA 94403(US)

(72) Inventor: Cunningham, Brian C.
24 Olive Avenue
Piedmont CA 94611(US)

(72) Inventor: Caldwell, Robert Mark
1828 Broadway
No.101 San Francisco Ca 94109(US)

(72) Inventor: Bott, Richard Ray
3032 Hillside drive
Burlingame CA 94010(US)

(72) Inventor: Estell, David Aaron
250 Diablo Avenue
Mountan View CA 94043(US)

(72) Inventor: Power, Scott Douglas
732 Olive Court
San Bruno CA 94066(US)

(74) Representative: Sibley, Richard Edward et al,
BOULT, WADE & TENNANT 27 Furnival Street
London EC4A 1PQ(GB)

(54) Non-human Carbonyl hydrolase mutants, DNA sequences and vectors encoding same and hosts transformed with said vectors.

(57) Novel carbonyl hydrolase mutants derived from the amino acid sequence of naturally-occurring or recombinant non-human carbonyl hydrolases and DNA sequences encoding the same. The mutant carbonyl hydrolases, in general, are obtained by *in vitro* modification of a precursor DNA sequence encoding the naturally-occurring or recombinant carbonyl hydrolase to encode the substitution, insertion or deletion of one or more amino acids in the amino acid sequence of a precursor carbonyl hydrolase. Such mutants have one or more properties which are different than the same property of the precursor hydrolase.

EP 0 251 446 A2

NON-HUMAN CARBONYL HYDROLASE MUTANTS, DNA SEQUENCES AND VECTORS ENCODING SAME AND HOSTS TRANSFORMED WITH SAID VECTORS

The recent development of various *in vitro* techniques to manipulate the DNA sequences encoding naturally-occuring polypeptides as well as recent developments in the chemical synthesis of relatively short sequences of single and double stranded DNA has resulted in the speculation that such techniques can be used to modify enzymes to improve some functional property in a predictable way. Ulmer, K.M. (1983) *Science* 219, 666-671. The only working example disclosed therein is the substitution of a single amino acid within the active site of tyrosyl-tRNA synthetase (Cys35-Ser) which lead to a reduction in enzymatic activity. See Winter, G., et al. (1982) *Nature* 299, 756-758; and Wilkinson, A.J., et al. (1983) *Biochemistry* 22, 3581-3586 (Cys35-Gly mutation also resulted in decreased activity).

When the same t-RNA synthetase was modified by substituting a different amino acid residue within the active site with two different amino acids, one of the mutants (Thr51-Ala) reportedly demonstrated a predicted moderate increase in kcat/Km whereas a second mutant (Thr51-Pro) demonstrated a massive increase in kcat/Km which could not be explained with

certainty. Wilkinson, A.H., et al. (1984) Nature 307, 187-188.

Another reported example of a single substitution of an amino acid residue is the substitution of cysteine for isoleucine at the third residue of T4 lysozyme. Perry, L.J., et al. (1984) Science 226, 555-557. The resultant mutant lysozyme was mildly oxidized to form a disulfide bond between the new cysteine residue at position 3 and the native cysteine at position 97. This crosslinked mutant was initially described by the author as being enzymatically identical to, but more thermally stable than, the wild type enzyme. However, in a "Note Added in Proof", the author indicated that the enhanced stability observed was probably due to a chemical modification of cysteine at residue 54 since the mutant lysozyme with a free thiol at Cys54 has a thermal stability identical to the wild type lysozyme.

Similarly, a modified dihydrofolate reductase from E.coli has been reported to be modified by similar methods to introduce a cysteine which could be crosslinked with a naturally-occurring cysteine in the reductase. Villafranca, D.E., et al. (1983) Science 222, 782-788. The author indicates that this mutant is fully reactive in the reduced state but has significantly diminished activity in the oxidized state. In addition, two other substitutions of specific amino acid residues are reported which resulted in mutants which had diminished or no activity.

EPO Publication No. 0130756 discloses the substitution of specific residues within B. amyloliquefaciens subtilisin with specific amino acids. Thus, Met222 has been substituted with all 19 other amino acids,

Gly166 with 9 different amino acids and Gly169 with Ala and Ser.

As set forth below, several laboratories have also reported the use of site directed mutagensis to produce the mutation of more than one amino acid residue within a polypeptide.

The amino-terminal region of the signal peptide of the prolipoprotein of the E. coli outer membrane was stated to be altered by the substitution or deletion of residues 2 and 3 to produce a charge change in that region of the polypeptide. Inoyye, S., et al. (1982) Proc. Nat. Acad. Sci. USA 79, 3438-3441. The same laboratory also reported the substitution and deletion of amino acid redisues 9 and 14 to determine the effects of such substitution on the hydrophobic region of the same signal sequence. Inouye, S., et al. (1984) J. Biol. Chem. 259, 3729-3733.

Double mutants in the active site of tyrosyl-t-RNA synthetase have also been reported. Carter, P.J., et al. (1984) Cell 38, 835-840. In this report, the improved affinity of the previously described Thr51-Pro mutant for ATP was probed by producing a second mutation in the active site of the enzyme. One of the double mutants, Gly35/Pro51, reportedly demonstrated an unexpected result in that it bound ATP in the transition state better than was expected from the two single mutants. Moreover, the author warns, at least for one double mutant, that it is not readily predictable how one substitution alters the effect caused by the other substitution and that care must be taken in interpreting such substitutions.

A mutant is disclosed in U.S. Patent No. 4,532,207, wherein a polyarginine tail was attached to the C-terminal residue of β-urogastrone by modifying the DNA sequence encoding the polypeptide. As disclosed, the polyarginine tail changed the electrophoretic mobility of the urogastrone-polyaginine hybrid permiting selective purification. The polyarginine was subsequently removed, according to the patentee, by a polyarginine specific exopeptidase to produce the purified urogastrone. Properly construed, this reference discloses hybrid polypeptides which do not constitute mutant polypeptides containing the substitution, insertion or deletion of one or more amino acids of a naturally occurring polypeptide.

Single and double mutants of rat pancreatic trypsin have also been reported. Craik, C.S., et al. (1985) Science 228, 291-297. As reported, glycine residues at positions 216 and 226 were replaced with alanine residues to produce three trypsin mutants (two single mutants and one double mutant). In the case of the single mutants, the authors stated expectation was to observe a differential effect on Km. They instead reported a change in specificity (kcat/Km) which was primarily the result of a decrease in kcat. In contrast, the double mutant reportedly demonstrated a differential increase in Km for lysyl and arginyl substrates as compared to wild type trypsin but had virtually no catalytic activity.

The references discussed above are provided solely for their disclosure prior to the filing date of the instant case, and nothing herein is to be construed as an admission that the inventors are not entitled to antedate such disclosure by virtue of prior invention or priority based on earlier filed applications.

Based on the above references, however, it is apparent that the modification of the amino acid sequence of wild type enzymes often results in the decrease or destruction of biological activity.

Accordingly, it is an object herein to provide carbonyl hydrolase mutants which have at least one property which is different from the same property of the carbonyl hydrolase precursor from which the amino acid of said mutant is derived.

It is a further object to provide mutant DNA sequences encoding such carbonyl hydrolase mutants as well as expression vectors containing such mutant DNA sequences.

Still further, another object of the present invention is to provide host cells transformed with such vectors as well as host cells which are capable of expressing such mutants either intracellularly or extracellularly.

Summary of the Invention

The invention includes carbonyl hydrolase mutants, preferably having at least one property which is substantially different from the same property of the precursor non-human carbonyl hydrolase from which the amino acid sequence of the mutant is derived. These properties include oxidative stability, substrate, specificity catalytic activity, thermal stability, alkaline stability, pH activity profile and resistance to proteolytic degradation. The precursor carbonyl hydrolase may be naturally occurring carbonyl hydrolases or recombinant carbonyl hydrolases. The amino acid sequence of the carbonyl hydrolase mutant is derived by the substitution, deletion or insertion of one or more amino acids of the precursor carbonyl hydrolase amino acid sequence.

The invention also includes mutant DNA sequences encoding such carbonyl hydrolase mutants. Further the invention includes expression vectors containing such mutant DNA sequences as well as host cells transformed with such vectors which are capable of expressing said carbonyl hydrolase mutants.

Brief Description of the Drawings

Figure 1 shows the nucleotide sequence of the coding strand, correlated with the amino acid sequence of B. amyloliquefaciens subtilisin gene. Promoter (P) ribosome binding site (rbs) and termination (term) regions of the DNA sequence as well as sequences encoding the presequence (PRE) putative prosequence (PRO) and mature form (MAT) of the hydrolase are also shown.

Figure 2 is a schematic diagram showing the substrate binding cleft of subtilisin together with substrate.

Figure 3 is a stereo view of the S-1 binding subsite of B. amyloliquefaciens subtilisin showing a lysine P-1 substrate bound in the site in two different ways. Figure 3A shows Lysine P-1 substrate bound to form a salt bridge with a Glu at position 156. Figure 3B shows Lysine P-1 substrate bound to form a salt bridge with Glu at position 166.

Figure 4 is a schematic diagram of the active site of subtilisin Asp32, His64 and Ser221.

Figures 5A and 5B depict the amino acid sequence of subtilisin obtained from various sources. The residues directly beneath each residue of B. amyloliquefaciens subtilisin are equivalent residues which (1) can be mutated in a similar manner to that described for B. amyloliquefaciens subtilisin, or (2) can be used as a replacement amino acid residue in B. amyloliquefaciens subtilisin. Figure 5C depicts conserved residues of B. amyloliquefaciens subtilisin when compared to other subtilisin sequences.

Figures 6A and 6B depict the inactivation of the mutants Met222L and Met222Q when exposed to various organic oxidants.

Figure 7 depicts the ultraviolet spectrum of Met222F subtilisin and the difference spectrum generated after inactivation by diperdodecanoic acid (DPDA).

Figure 8 shows the pattern of cyanogen bromide digests of untreated and DPDA oxidized subtilisin Met222F on high resolution SDS-pyridine peptide gels.

Figure 9 depicts a map of the cyanogen bromide fragments of Fig. 8 and their alignment with the sequence of subtilisin Met222F.

Figure 10 depicts the construction of mutations between codons 45 and 50 of B. amyloliquefaciens subtilisin.

Figure 11 depicts the construction of mutations between codons 122 and 127 of B. amyloliquefaciens subtilisin.

Figure 12 depicts the effect of DPDA on the activity of subtilisin mutants at positions 50 and 124 in subtilisin Met222F.

Figure 13 depicts the construction of mutations at codon 166 of B. amyloliquefaciens subtilisin.

Figure 14 depicts the effect of hydrophobicity of the P-1 substrate side-chain on the kinetic parameters of wild-type B. amyloliquefaciens subtilisin.

Figure 15 depicts the effect of position 166 side-chain substitutions on P-1 substrate specificity. Figure 15A shows position 166 mutant subtilisins containing non-branched alkyl and aromatic side-chain substitutions arranged in order of increasing molecular volume. Figure 15B shows a series of mutant enzymes progressing through β- and γ-branched aliphatic side chain substitutions of increasing molecular volume.

Figure 16 depicts the effect of position 166 side-chain volumn on log kcat/Km for various P-1 substrates.

Figure 17 shows the substrate specificity differences between Ile166 and wild-type (Gly166) B. amyloliquefaciens subtilisin against a series of alphatic and aromatic substrates. Each bar represents the difference in log kcat/Km for Ile166 minus wild-type (Gly166) subtilisin.

Figure 18 depicts the construction of mutations at codon 169 of B. amyloliquefaciens subtilisin.

Figure 19 depicts the construction of mutations at codon 104 of B. amyloliquefaciens subtilisin.

Figure 20 depicts the construction of mutations at codon 152 B. amyloliquefaciens subtilisin.

Figure 21 depicts the construction of single mutations at codon 156 and double mutations at codons 156 and 166 of B. amyloliquefaciens subtilisin.

Figure 22 depicts the construction of mutations at codon 217 for B. amyloliquefaciens subtilisin.

Figure 23 depicts the kcat/Km versus pH profile for mutations at codon 156 and 166 in B. amyloliquefaciens subtilisin.

Figure 23A depicts the kcat/Km versus pH profile for mutations at codon 156 and 166 in B. amyloliquefaciens subtilisin.

Figure 24 depicts the kcat/Km versus pH profile for mutations at codon 222 in B. amyloliquefaciens subtilisin.

Figure 25 depicts the constructing mutants at codons 94, 95 and 96.

Figures 26 and 27 depict substrate specificity of various wild type and mutant subtilisins for different substrates.

Figures 28 A, B, C and D depict the effect of charge in the P-1 binding sites due to substitutions at codon 156 and 166.

Figures 29 A and B are a stereoview of the P-1 binding site of subtilisin BPN' showing a lysine P-1 substrate bound in the site in two ways. In 29A, Lysine P-1 substrate is built to form a salt bridge with a Glu at codon 156. In 29B, Lysine P-1 substrate is built to form a salt bridge with Glu at codon 166.

Figure 30 demonstrates residual enzyme activity versus temperature curves for purified wild-type (Panel A), C22/C87 (Panel B) and C24/C87 (Panel C).

Figure 31 depicts the strategy for producing point mutations in the subtilisin coding sequence by misincorporation of $^{\alpha}$-thioldeoxynucleotide triphosphates.

Figure 32 depicts the autolytic stability of purified wild type and mutant subtilisins 170E, 107V, 213R and 107V/213R at alkaline pH.

Figure 33 depicts the autolytic stability of purified wild type and mutant subtilisins V50, F50 and F50/V107/R213 at alkaline pH.

Figure 34 depicts the strategy for constructing plasmids containing random cassette mutagenesis over residues 197 through 228.

Figure 35 depicts the oligodeoxynucleotides used for random cassette mutagenesis over residues 197 through 228.

Figure 36 depicts the construction of mutants at codon 204.

Figure 37 depicts the oligodeoxynucleotides used for synthesizing mutants at codon 204.

Detailed Description

The inventors have discovered that various single and multiple *in vitro* mutations involving the substitution, deletion or insertion of one or more amino acids within a non-human carbonyl hydrolase amino acid sequence can confer advantageous properties to such mutants when compared to the non-mutated carbonyl hydrolase.

Specifically, *B. amyloliquefaciens* subtilisin, an alkaline bacterial protease, has been mutated by modifying the DNA encoding the subtilisin to encode the substitution of one or more amino acids at various amino acid residues within the mature form of the subtilisin molecule. These *in vitro* mutant subtilisins have at least one property which is different when compared to the same property of the precursor subtilisin. These modified properties fall into several categories including: oxidative stability, substrate specificity, thermal stability, alkaline stability, catalytic activity, pH activity

profile, resistance to proteolytic degradation, Km, kcat and Km/kcat ratio.

Carbonyl hydrolases are enzymes which hydrolyze

compounds containing C-X bonds in which X is oxygen or nitrogen. They include naturally-occurring carbonyl hydrolases and recombinant carbonyl hydrolases. Naturally occurring carbonyl hydrolases principally include hydrolases, e.g. lipases and peptide hydrolases, e.g. subtilisins or metalloproteases. Peptide hydrolases include α-aminoacylpeptide hydrolase, peptidylamino-acid hydrolase, acylamino hydrolase, serine carboxypeptidase, metallocarboxy-peptidase, thiol proteinase, carboxylproteinase and metalloproteinase. Serine, metallo, thiol and acid proteases are included, as well as endo and exo-proteases.

"Recombinant carbonyl hydrolase" refers to a carbonyl hydrolase in which the DNA sequence encoding the naturally occurring carbonyl hydrolase is modified to produce a mutant DNA sequence which encodes the substitution, insertion or deletion of one or more amino acids in the carbonyl hydrolase amino acid sequence. Suitable modification methods are disclosed herein and in EPO Publication No. 0130756 published January 9, 1985.

Subtilisins are bacterial carbonyl hydrolases which generally act to cleave peptide bonds of proteins or peptides. As used herein, "subtilisin" means a naturally occurring subtilisin or a recombinant subtilisin. A series of naturally occurring subtilisins is known to be produced and often secreted

by various bacterial species. Amino acid sequences of the members of this series are not entirely homologous. However, the subtilisins in this series exhibit the same or similar type of proteolytic activity. This class of serine proteases shares a common amino acid sequence defining a catalytic triad which distinguishes them from the chymotrypsin related class of serine proteases. The subtilisins and chymotrypsin related serine proteases both have a catalytic triad comprising aspartate, histidine and serine. In the subtilisin related proteases the relative order of these amino acids, reading from the amino to carboxy terminus is aspartate-histidine-serine. In the chymotrypsin related proteases tne relative order, however is histidine-aspartate-serine. Thus, subtilisin herein refers to a serine protease having the catalytic triad of subtilisin related proteases.

"Recombinant subtilisin" refers to a subtilisin in which the DNA sequence encoding the subtilisin is modified to produce a mutant DNA sequence which encodes the substitution, deletion or insertion of one or more amino acids in the naturally occurring subtilisin amino acid sequence. Suitable methods to produce such modification include those disclosed herein and in EPO Publication No. 0130756. For example, the subtilisin multiple mutant herein containing the substitution of methionine at amino acid residues 50, 124 and 222 with phenylalanine, isoleucine and glutamine, respectively, can be considered to be derived from the recombinant subtilisin containing the substitution of glutamine at residue 222 (Q222) disclosed in EPO Publication No. 0130756. The multiple mutant thus is produced by the substitution of phenylalanine for methionine at

residue 50 and isoleucine for methionine at residue 124 in the Q222 recombinant subtilisin.

"Carbonyl hydrolases" and their genes may be obtained from many procaryotic and eucaryotic organisms. Suitable examples of procaryotic organisms include gram negative organisms such as E. coli or pseudomonas and gram positive bacteria such as micrococcus or bacillus. Examples of eucaryotic organisms from which carbonyl hydrolase and their genes may be obtained include yeast such as S. cerevisiae, fungi such as Aspergillus sp., and non-human mammalian sources such as, for example, Bovine sp. from which the gene encoding the carbonyl hydrolase chymosin can be obtained. As with subtilisins, a series of carbonyl hydrolases can be obtained from various related species which have amino acid sequences which are not entirely homologous between the members of that series but which nevertheless exhibit the same or similar type of biological activity. Thus, non-human carbonyl hydrolase as used herein has a functional definition which refers to carbonyl hydrolases which are associated, directly or indirectly, with procaryotic and non-human eucaryotic sources.

A "carbonyl hydrolase mutant" has an amino acid sequence which is derived from the amino acid sequence of a non-human "precursor carbonyl hydrolase". The precursor carbonyl hydrolases include naturally-occurring carbonyl hydrolases and recombinant carbonyl hydrolases. The amino acid sequence of the carbonyl hydrolase mutant is "derived" from the precursor hydrolase amino acid sequence by the substitution, deletion or insertion of one or more amino acids of the precursor amino acid sequence. Such modification is of the "precursor DNA sequence" which encodes the

amino acid sequence of the precursor carbonyl hydrolase rathern than manipulation of the precursor carbonyl hydrolase per se. Suitable methods for such manipulation of the precursor DNA sequence include methods disclosed herein and in EPO Publication No. 0130756.

Specific residues of B. amyloliquefaciens subtilisin are identified for substitution, insertion or deletion. These amino acid position numbers refer to those assigned to the B. amyloliquefaciens subtilisin sequence presented in Fig. 1. The invention, however, is not limited to the mutation of this particular subtilisin but extends to precursor carbonyl hydrolases containing amino acid residues which are "equivalent" to the particular identified residues in B. amyloliquefaciens subtilisin.

A residue (amino acid) of a precursor carbonyl hydrolase is equivalent to a residue of B. amyloliquefaciens subtilisin if it is either homologous (i.e., corresponding in position in either primary or tertiary structure) or analagous to a specific residue or portion of that residue in B. amyloliquefaciens subtilisin (i.e., having the same or similar functional capacity to combine, react, or interact chemically).

In order to establish homology to primary structure, the amino acid sequence of a precursor carbonyl hydrolase is directly comparted to the B. amyloliquefaciens subtilisin primary sequence and particularly to a set of residues known to be invariant in all subtilisins for which sequence is known (Figure 5C). After aligning the conserved residues, allowing for necessary insertions and

deletions in order to maintain alignment (i.e., avoiding the elimination of conserved residues through arbitrary deletion and insertion), the residues equivalent to particular amino acids in the primary sequence of B. amyloliquefaciens subtilisin are defined. Alignment of conserved residues preferably should conserve 100% of such residues. However, alignment of greater than 75% or as little as 50% of conserved residues is also adequate to define equivalent residues. Conservation of the catalytic triad, Asp32/His64/Ser221 should be maintained.

For example, in Figure 5A the amino acid sequence of subtilisin from B. amyloliquefaciens B. subtilisin var. I168 and B. lichenformis (carlsbergensis) are aligned to provide the maximum amount of homology between amino acid sequences. A comparison of these sequences shows that there are a number of conserved residues contained in each sequence. These residues are identified in Fig. 5C.

These conserved residues thus may be used to define the corresponding equivalent amino acid residues of B. amyloliquefaciens subtilisin in other carbonyl hydrolases such as thermitase derived from Thermoactinomyces. These two particular sequences are aligned in Fig. 5B to produce the maximum homology of conserved residues. As can be seen there are a number of insertions and deletions in the thermitase sequence as compared to B. amyloliquefaciens subtilisin. Thus, in thermitase the equivalent amino acid of Tyr217 in B. amyloliquefaciens subtilisin is the particular lysine shown beneath Tyr217.

In Fig. 5A, the equivalent amino acid at position 217 in B. amyloliquefaciens subtilisin is Tyr. Likewise,

in B. subtilis subtilisin position 217 is also occupied by Tyr but in B. licheniformis position 217 is occupied by Leu.

Thus, these particular residues in thermitase, and subtilisin from B. subtilisin and B. licheniformis may be substituted by a different amino acid to produce a mutant carbonyl hydrolase since they are equivalent in primary structure to Tyr217 in B. amyloliquefaciens subtilisin. Equivalent amino acids of course are not limited to those for Tyr217 but extend to any residue which is equivalent to a residue in B. amyloliquefaciens whether such residues are conserved or not.

Equivalent residues homologous at the level of tertiary structure for a precursor carbonyl hydrolase whose tertiary structure has been determined by x-ray crystallography, are defined as those for which the atomic coordinates of 2 or more of the main chain atoms of a particular amino acid residue of the precursor carbonyl hydrolase and B. amyloliquefaciens subtilisin (N on N, CA on CA, C on C, and O on O) are within 0.13nm and preferably 0.1nm after alignment. Alignment is achieved after the best model has been oriented and positioned to give the maximum overlap of atomic coordinates of non-hydrogen protein atoms of the carbonyl hydrolase in question to the B. amyloliquefaciens subtilisin. The best model is the crystallographic model giving the lowest R factor for experimental diffraction data at the highest resolution available.

$$\text{R factor} = \frac{\sum_h |Fo(h)| - |Fc(h)|}{\sum_h |Fo(h)|}$$

Equivalent residues which are functionally analogous to a specific residue of B. amyloliquefaciens subtilisin are defined as those amino acids of the precursor carbonyl hydrolases which may adopt a conformation such that they either alter, modify or contribute to protein structure, substrate binding or catalysis in a manner defined and attributed to a specific residue of the B. amyloliquefaciens subtilisin as described herein. Further, they are those residues of the precursor carbonyl hydrolase (for which a tertiary structure has been obtained by x-ray crystallography), which occupy an analogous position to the extent that although the main chain atoms of the given residue may not satisfy the criteria of equivalence on the basis of occupying a homologous position, the atomic coordinates of at least two of the side chain atoms of the residue lie with 0.13nm of the corresponding side chain atoms of B. amyloliquefaciens subtilisin. The three dimensional structures would be aligned as outlined above.

Some of the residues identified for substitution, insertion or deletion are conserved residues whereas others are not. In the case of residues which are not conserved, the replacement of one or more amino acids is limited to substitutions which produce a mutant which has an amino acid sequence that does not correspond to one found in nature. In the case of conserved residues, such replacements should not result in a naturally occurring sequence. The carbonyl hydrolase mutants of the present invention include the mature forms of carbonyl hydrolase mutants as well as the pro- and prepro-forms of such hydrolase mutants. The prepro-forms are the preferred construction since

this facilitates the expression, secretion and maturation of the carbonyl hydrolase mutants.

"Expression vector" refers to a DNA construct containing a DNA sequence which is operably linked to a suitable control sequence capable of effecting the expression of said DNA in a suitable host. Such control sequences include a promoter to effect transcription, an optional operator sequence to control such transcription, a sequence encoding suitable mRNA ribosome binding sites, and sequences which control termination of transcription and translation. The vector may be a plasmid, a phage particle, or simply a potential genomic insert. Once transformed into a suitable host, the vector may replicate and function independently of the host genome, or may, in some instances, integrate into the genome itself. In the present specification, "plasmid" and "vector" are sometimes used interchangeably as the plasmid is the most commonly used form of vector at present. However, the invention is intended to include such other forms of expression vectors which serve equivalent functions and which are, or become, known in the art.

The "host cells" used in the present invention generally are procaryotic or eucaryotic hosts which preferably have been manipulated by the methods disclosed in EPO Publication No. 0130756 to render

them incapable of secreting enzymatically active endoprotease. A preferred host cell for expressing subtilisin is the Bacillus strain BG2036 which is deficient in enzymatically active neutral protease and alkaline protease (subtilisin). The construction of strain BG2036 is described in detail in EPO Publicatin No. 0130756 and further described by Yang, M.Y., et al. (1984) J. Bacteriol. 160, 15-21. Other host cells for expressing subtilisin include Bacillus subtilis I168 (EPO Publication No. 0130756).

Host cells are transformed or transfected with vectors constructed using recombinant DNA techniques. Such transformed host cells are capable of either replicating vectors encoding the carbonyl hydrolase mutants or expressing the desired carbonyl hydrolase mutant. In the case of vectors which encode the pre or prepro form of the carbonyl hydrolase mutant, such mutants, when expressed, are typically secreted from the host cell into the host cell medium.

"Operably linked" when describing the relationship between two DNA regions simply means that they are functionally related to each other. For example, a presequence is operably linked to a peptide if it functions as a signal sequence, participating in the secretion of the mature form of the protein most probably involving cleavage of the signal sequence. A promoter is operably linked to a coding sequence if it controls the transcription of the sequence; a ribosome binding site is operably linked to a coding sequence if it is positioned so as to permit translation.

The genes encoding the naturally-occurring precursor carbonyl hydrolase may be obtained in accord with the

general methods described herein in EPO Publication No. 0130756.

Once the carbonyl hydrolase gene has been cloned, a number of modifications are undertaken to enhance the use of the gene beyond synthesis of the naturally-occurring precursor carbonyl hydrolase. Such modifications include the production of recombinant carbonyl hydrolases as disclosed in EPO Publication No. 0130756 and the production of carbonyl hydrolase mutants described herein.

The carbonyl hydrolase mutants of the present invention may be generated by site specific mutagenesis (Smith, M. (1985) Ann, Rev. Genet. 423; Zoeller, M.J., et al. (1982) Nucleic Acid Res. 10, 6487-6500), cassette mutagenesis (EPO Publication No. 0130756) or random mutagenesis (Shortle, D., et al. (1985) Genetics, 110, 539; Shortle, D., et al. (1986) Proteins: Structure, Function and Genetics, 1, 81; Shortle, D. (1986) J. Cell. Biochem, 30, 281; Alber, T., et al. (1985) Proc. Natl. Acad. of Sci., 82, 747; Matsumura, M., et al. (1985) J. Biochem., 260, 15298; Liao, H., et al. (1986) Proc. Natl. Acad. of Sci., 83 576) of the cloned precursor carbonyl hydrolase. Cassette mutagenesis and the random mutagenesis method disclosed herein are preferred.

The mutant carbonyl hydrolases expressed upon transformation of suitable hosts are screened for enzymes exhibiting one or more properties which are substantially different from the properties of the precursor carbonyl hydrolases, e.g., changes in substrate specificity, oxidative stability, thermal stability, alkaline stability, resistance to

proteolytic degradation, pH-activity profiles and the like.

A change in substrate specificity is defined as a difference between the kcat/Km ratio of the precursor carbonyl hydrolase and that of the hydrolase mutant. The kcat/Km ratio is a measure of catalytic efficienty. Carbonyl hydrolase mutants with increased or diminished kcat/Km ratios are described in the examples. Generally, the objective will be to secure a mutant having a greater (numerically large) kcat/Km ratio for a given substrate, thereby enabling the use of the enzyme to more efficiently act on a target substrate. A substantial change in kcat/Km ratio is preferably at least 2-fold increase or decrease. However, smaller increases or decreases in the ratio (e.g., at least 1.5-fold) are also considered substantial. An increase in kcat/Km ratio for one substrate may be accompanied by a reduction in kcat/Km ratio for another substrate. This is a shift in substrate specificity, and mutants exhibiting such shifts have utility where the precursor hydrolase is undesirable, e.g. to prevent undesired hydrolysis of a particular substrate in an admixture of substrates. Km and kcat are measured in accord with known procedures, as described in EPO Publication No. 0130756 or as described herein.

Oxidative stability is measured either by known procedures or by the methods described hereinafter. A substantial change in oxidative stability is evidenced by at least about 50% increase or decrease (preferably decrease) in the rate of loss of enzyme activity when exposed to various oxidizing conditions. Such oxidizing conditions are exposure to the organic

oxidant diperdodecanoic acid (DPDA) under the conditions described in the examples.

Alkaline stability is measured either by known procedures or by the methods described herein. A substantial change in alkaline stability is evidenced by at least about a 5% or greater increase or decrease (preferably increase) in the half life of the enzymatic activity of a mutant when compared to the precursor carbonyl hydrolase. In the case of subtilisins, alkaline stability was measured as a function of autoproteolytic degradation of subtilisin at alkaline pH, e.g. for example, 0.1M sodium phosphate, pH 12 at 25° or 30°C.

Thermal stability is measured either by known procedures or by the methods described herein. A substantial change in thermal stability is evidenced by at least about a 5% or greater increase or decrease (preferably increase) in the half-life of the catalytic activity of a mutant when exposed to a relatively high temperature and neutral pH as compared to the precursor carbonyl hydrolase. In the case of subtilisins, thermal stability is measured by the autoproteolytic degradation of subtilisin at elevated temperatures and neutral pH, e.g., for example 2mM calcium chloride, 50mM MOPS pH 7.0 at 59°C.

The inventors have produced mutant subtilisins containing the substitution of the amino acid residues of B. amyloliquefaciens subtilisin shown in Table I. The wild type amino acid sequence and DNA sequence of B. amyloliquefaciens subtilisin is shown in Fig. 1.

TABLE I

Residue	Replacement Amino Acid
Tyr21	F A
Thr22	C
Ser24	C
Asp32	Q S
Ser33	A T
Asp36	A G
Gly46	V
Ala48	E V R
Ser49	C L
Met50	C F V
Asn77	D
Ser87	C
Lys94	C
Val95	C
Leu96	D
Tyr104	A C D E F G H I K L M N P Q R S T V W
Ile107	V
Gly110	C R
Met124	I L
Asn155	A D H Q T
Glu156	Q S
Gly166	C E I L M P S T W Y
Gly169	C D E F H I K L M N P Q R T V W Y
Lys170	E R
Tyr171	F
Pro172	E Q
Phe189	A C D E G H I K L M N P Q R S T V W Y
Asp197	R A
Met199	I
Ser204	C R L P
Lys213	R T
Tyr217	A C D E F G H I K L M N P Q R S T V W
Ser221	A C

The different amino acids substituted are represented in Table I by the following single letter designations:

Amino acid or residue thereof	3-letter symbol	1-letter symbol
Alanine	Ala	A
Glutamate	Glu	E
Glutamine	Gln	Q
Aspartate	Asp	D
Asparagine	Asn	N
Leucine	Leu	L
Glycine	Gly	G
Lysine	Lys	K
Serine	Ser	S
Valine	Val	V
Arginine	Arg	R
Threonine	Thr	T
Proline	Pro	P
Isoleucine	Ile	I
Methionine	Met	M
Phenylalanine	Phe	F
Tyrosine	Tyr	Y
Cysteine	Cys	C
Tryptophan	Trp	W
Histidine	His	H

Except where otherwise indicated by context, wild-type amino acids are represented by the above three-letter symbols and replaced amino acids by the above single-letter symbols. Thus, if the methionine at residue 50 in B. amyloliquefaciens subtilisin is

replaced by phenylalanine, this mutation (mutant) may be designated Met50F or F50. Similar designations are used for multiple mutants.

In addition to the amino acids used to replace the residues disclosed in Table I, other replacements of amino acids at these residues are expected to produce mutant subtilisins having useful properties. These residues and replacement amino acids are shown in Table II.

TABLE II

Residue	Replacement Amino Acid(s)
Tyr-21	L
Thr22	K
Ser24	A
Asp32	
Ser33	G
Gly46	
Ala48	
Ser49	
Met50	L K I V
Asn77	D
Ser87	N
Lys94	R Q
Val95	L I
Tyr104	
Met124	K A
Ala152	C L I T M
Asn155	
Glu156	A T M L V
Gly166	
Gly169	
Tyr171	K R E Q
Pro172	D N
Phe189	
Tyr217	
Ser221	
Met222	

Each of the mutant subtilisins in Table I contain the replacement of a single residue of the B. amyloliquefaciens amino acid sequence. These particular residues were chosen to probe the influence

of such substitutions on various properties of B. amyloliquefacien subtilisin.

Thus, the inventors have identified Met124 and Met222 as important residues which if substituted with another amino acid produce a mutant subtilisin with enhanced oxidative stability. For Met124, Leu and Ile are preferred replacement amino acids. Preferred amino acids for replacement of Met222 are disclosed in EPO Publication No. 0130756.

Various other specific residues have also been identified as being important with regard to substrate specificity. These residues include Tyr104, Ala152, Glu156, Gly166, Gly169, Phe189 and Tyr217 for which mutants containing the various replacement amino acids presented in Table I have already been made, as well as other residues presented below for which mutants have yet to be made.

The identification of these residues, including those yet to be mutated, is based on the inventors' high resolution crystal structure of B. amyloliquefaciens subtilisin to 1.8 A (see Table III), their experience with in vitro mutagenesis of subtilisin and the literature on subtilisin. This work and the x-ray crystal structures of subtilisin containing covalently bound peptide inhibitors (Robertus, J.D., et al. (1972) Biochemistry 11, 2439-2449), product complexes (Robertus, J.D., et al. (1972) Biochemistry 11, 4293-4303), and transition state analogs (Matthews, D.A., et al (1975) J. Biol. Chem. 250, 7120-7126; Poulos, T.L., et al. (1976) J. Biol. Chem. 251, 1097-1103), has helped in identifying an extended peptide binding cleft in subtilisin. This substrate binding cleft together with substrate is schematically

diagramemed in Fig. 2, according to the nomenclature of Schechter, I., et al. (1967) Biochem Bio. Res. Commun. 27, 157. The scissile bond in the substrate is identified by an arrow. The P and P' designations refer to the amino acids which are positioned respectively toward the amino or carboxy terminus relative to the scissle bond. The S and S' designations refer to subsites in the substrate binding cleft of subtilisin which interact with the corresponding substrate amino acid residues.

Atomic Coordinates for the
Apoenzyme Form of B, Amyloliquefaciens
Subtilisin to 1.8AResolution

1	ALA	N	19.436	53.195	-21.756	1	ALA	CA	19.811	51.774	-21.965
1	ALA	C	18.731	50.825	-21.324	1	ALA	O	18.376	51.197	-20.175
1	ALA	CB	21.099	51.318	-21.183	2	GLN	N	18.268	49.884	-22.841
2	GLN	CA	17.219	49.808	-21.434	2	GLN	C	17.875	47.704	-20.992
2	GLN	O	18.765	47.165	-21.891	2	GLN	CB	16.125	48.768	-22.449
2	GLN	CG	15.328	47.805	-21.927	2	GLN	CD	13.912	47.742	-22.936
2	GLN	OE1	13.823	48.612	-22.867	2	GLN	NE2	14.115	46.917	-23.926
3	SER	N	17.477	47.285	-19.852	3	SER	CA	17.950	45.868	-19.437
3	SER	C	16.735	44.918	-19.469	3	SER	O	15.590	45.352	-19.229
3	SER	CB	18.588	45.838	-18.869	3	SER	OG	17.882	46.218	-17.869
4	VAL	N	16.991	43.646	-19.725	4	VAL	CA	15.966	42.619	-19.639
4	VAL	C	16.129	41.934	-18.298	4	VAL	O	17.123	41.178	-18.886
4	VAL	CB	16.808	41.622	-20.822	4	VAL	CG1	14.874	48.572	-20.741
4	VAL	CG2	16.837	42.266	-22.186	5	PRO	N	15.239	42.306	-17.331
5	PRO	CA	15.384	41.415	-16.827	5	PRO	C	15.581	39.955	-16.249
5	PRO	O	16.885	39.263	-17.146	5	PRO	CB	14.158	41.880	-15.283
5	PRO	CG	13.841	43.215	-15.921	5	PRO	CD	14.864	42.986	-17.417
6	TYR	N	16.363	39.240	-15.487	6	TYR	CA	16.628	37.883	-15.715
6	TYR	C	15.359	36.975	-15.528	6	TYR	O	15.224	35.943	-16.235
6	TYR	CB	17.824	37.323	-14.834	6	TYR	CG	16.821	35.847	-15.855
6	TYR	CD1	18.437	35.452	-16.346	6	TYR	CD2	17.696	34.988	-14.871
6	TYR	CE1	18.535	34.870	-16.653	6	TYR	CE2	17.815	33.539	-14.379
6	TYR	CZ	18.222	33.154	-15.628	6	TYR	OH	18.312	31.838	-15.998
7	GLY	N	14.666	37.362	-14.638	7	GLY	CA	13.211	36.668	-14.376
7	GLY	C	12.488	36.535	-15.478	7	GLY	O	11.747	35.478	-15.883
8	VAL	N	12.641	37.529	-16.541	8	VAL	CA	11.777	37.523	-17.836
8	VAL	C	12.363	36.433	-18.735	8	VAL	O	11.639	35.716	-18.478
8	VAL	CB	11.785	38.908	-18.567	8	VAL	CG1	12.196	38.883	-19.963
8	VAL	CG2	10.991	39.918	-17.733	9	SER	N	13.661	36.318	-18.775
9	SER	CA	14.419	35.342	-19.562	9	SER	C	14.188	33.920	-18.945
9	SER	O	14.112	33.814	-19.801	9	SER	CB	15.926	35.632	-19.585
9	SER	OG	16.182	36.747	-20.358	10	GLN	N	14.115	33.887	-17.862
10	GLN	CA	13.964	32.836	-16.876	10	GLN	C	12.887	31.887	-17.277
10	GLN	O	12.785	30.462	-17.413	10	GLN	CB	14.125	32.885	-15.418
10	GLN	CG	14.295	31.417	-14.588	10	GLN	CD	14.486	31.911	-13.167
10	GLN	OE1	14.554	33.848	-12.744	10	GLN	NE2	14.552	30.960	-12.251
11	ILE	N	11.625	32.575	-17.870	11	ILE	CA	10.373	31.904	-18.182
11	ILE	C	10.209	31.792	-19.605	11	ILE	O	9.173	31.333	-20.180
11	ILE	CB	9.132	32.669	-17.675	11	ILE	CG1	9.864	34.117	-18.849
11	ILE	CG2	8.142	32.455	-15.941	11	ILE	CD1	7.588	34.648	-17.923
12	LYS	N	11.272	32.385	-20.277	12	LYS	CA	11.388	32.119	-21.722
12	LYS	C	10.456	33.886	-22.522	12	LYS	O	10.178	32.783	-23.486
12	LYS	CB	11.257	36.446	-22.216	12	LYS	CG	12.283	29.830	-21.423
12	LYS	CD	12.543	28.517	-22.159	12	LYS	CE	13.823	27.467	-21.166
12	LYS	NZ	14.476	27.480	-20.935	13	ALA	N	10.189	34.138	-21.991
13	ALA	CA	9.323	35.198	-22.631	13	ALA	C	10.826	35.716	-23.843
13	ALA	O	9.338	35.804	-24.901	13	ALA	CB	8.885	36.195	-21.565
14	PRO	N	11.332	35.958	-23.893	14	PRO	CA	11.995	36.439	-25.128
14	PRO	C	11.786	35.557	-26.317	14	PRO	O	11.778	36.847	-27.445
14	PRO	CB	13.462	36.588	-24.692	14	PRO	CG	13.328	36.978	-23.221
14	PRO	CD	12.281	35.936	-22.758	15	ALA	N	11.648	34.236	-26.129
15	ALA	CA	11.379	33.458	-27.367	15	ALA	C	10.882	33.795	-28.832
15	ALA	O	10.888	33.718	-29.278	15	ALA	CB	11.952	31.949	-27.862
16	LEU	N	9.883	34.138	-27.248	16	LEU	CA	9.791	34.558	-27.828
16	LEU	C	7.912	35.925	-28.521	16	LEU	O	9.342	36.124	-29.868
16	LEU	CB	8.746	34.623	-28.698	16	LEU	CG	8.798	33.443	-26.822
16	LEU	CD1	8.881	33.234	-27.809	16	LEU	CD2	6.494	32.287	-26.282
17	MIS	N	8.665	36.829	-27.922	17	MIS	CA	8.098	38.351	-28.538
17	MIS	C	9.318	37.981	-29.898	17	MIS	N	9.187	38.422	-30.856
17	MIS	CB	9.788	39.188	-27.652	17	MIS	CG	9.185	39.288	-26.262
17	MIS	SD1	9.938	39.887	-26.272	17	MIS	CE2	8.886	38.924	-25.494
17	MIS	CE1	9.224	39.934	-24.146	17	MIS	NE2	8.879	39.328	-24.381
18	SER	N	10.463	37.833	-30.822	18	SER	CA	11.189	36.739	-31.322

34	ASP	O	3.884	55.471	-13.579	
34	ASP	CG	4.339	57.899	-10.806	
34	ASP	OD2	3.468	57.277	-10.263	
37	SER	CA	1.183	57.221	-14.512	
37	SER	O	2.345	58.303	-14.151	
37	SER	OG	-0.860	59.133	-13.879	
38	SER	CA	0.261	59.505	-14.687	
38	SER	O	0.543	59.251	-15.285	
38	SER	OG	0.376	59.865	-12.234	
39	HIS	CA	0.637	56.574	-15.291	
39	HIS	O	0.738	55.878	-17.419	
39	HIS	CG	0.814	54.409	-14.456	
39	HIS	CD2	0.749	54.345	-13.389	
39	HIS	NE2	0.984	53.918	-13.808	
40	PRO	CA	7.988	56.897	-18.831	
40	PRO	O	8.832	55.897	-20.578	
40	PRO	CG	10.853	57.685	-17.982	
41	ASP	N	8.681	54.328	-18.485	
41	ASP	OD1	10.325	51.395	-20.429	
41	ASP	CB	9.799	52.239	-18.224	
41	ASP	C	7.311	52.163	-18.839	
42	LEU	N	6.185	52.803	-18.558	
42	LEU	C	3.924	52.987	-19.376	
42	LEU	CB	4.421	52.158	-17.808	
42	LEU	CD1	4.535	51.546	-16.581	
43	LYS	N	3.818	52.135	-19.944	
43	LYS	C	0.637	52.156	-20.818	
43	LYS	CB	2.822	52.388	-22.169	
43	LYS	CD	0.958	52.862	-24.339	
43	LYS	NZ	0.337	51.757	-26.418	
44	VAL	CA	-1.607	52.638	-18.765	
44	VAL	O	-2.623	53.986	-20.434	
44	VAL	CG1	-2.724	52.941	-16.582	
45	ALA	N	-3.494	51.931	-19.871	
45	ALA	C	-5.841	52.507	-20.053	
45	ALA	CB	-4.831	50.580	-21.389	
46	GLY	CA	-7.882	52.837	-18.681	
46	GLY	O	-5.938	52.806	-16.835	
47	GLY	CA	-8.014	52.248	-14.388	
47	GLY	O	-9.988	53.481	-14.185	
48	ALA	CA	-10.255	52.878	-11.382	
48	ALA	O	-9.864	51.728	-9.725	
49	SER	N	-10.149	53.547	-9.837	
49	SER	C	-10.947	52.986	-7.783	
49	SER	CB	-9.892	54.588	-7.029	
50	MET	N	-10.835	52.887	-5.932	
50	MET	C	-11.463	51.962	-3.561	
50	MET	CB	-12.812	50.818	-4.998	
50	MET	SD	-13.448	49.889	-7.256	
51	VAL	N	-10.427	52.748	-3.422	
51	VAL	C	-10.830	54.562	-1.987	
51	VAL	CB	-8.443	53.155	-2.988	
51	VAL	CG2	-7.784	51.815	-2.382	
52	PRO	CA	-12.372	55.933	-0.821	
52	PRO	O	-11.771	56.238	-0.925	
52	PRO	CG	-13.583	54.183	0.885	
53	SER	N	-10.442	56.886	0.299	
53	SER	C	-8.428	58.245	-0.326	
53	SER	CB	-9.884	57.787	2.049	
54	GLU	N	-8.254	57.523	-1.393	
54	GLU	C	-7.767	57.383	-3.785	
54	GLU	CB	-6.154	58.599	-2.154	
54	GLU	CG	-6.844	59.849	-0.878	
54	GLU	CD	-5.912	60.911	-0.348	
54	GLU	OE1	-4.691	60.722	-0.488	
54	GLU	OE2	-6.412	61.988	0.048	
55	ASP	N	-8.468	59.511	-4.588	
55	ASP	C	-8.888	61.888	-5.888	
55	ASP	CB	-8.888	61.888	-5.888	
55	ASP	CG	-8.888	61.888	-5.888	
55	ASP	OD1	-8.888	61.888	-5.888	
55	ASP	OD2	-8.888	61.888	-5.888	
56	ASP	N	-8.888	61.888	-5.888	
56	ASP	C	-8.888	61.888	-5.888	
56	ASP	CB	-8.888	61.888	-5.888	
56	ASP	CG	-8.888	61.888	-5.888	
56	ASP	OD1	-8.888	61.888	-5.888	
56	ASP	OD2	-8.888	61.888	-5.888	
57	SER	N	-8.888	61.888	-5.888	
57	SER	C	-8.888	61.888	-5.888	
57	SER	CB	-8.888	61.888	-5.888	
57	SER	OG	-8.888	61.888	-5.888	
58	MET	N	-8.888	61.888	-5.888	
58	MET	C	-8.888	61.888	-5.888	
58	MET	CB	-8.888	61.888	-5.888	
58	MET	CG	-8.888	61.888	-5.888	
58	MET	SD	-8.888	61.888	-5.888	
58	MET	CE	-8.888	61.888	-5.888	
59	VAL	N	-8.888	61.888	-5.888	
59	VAL	C	-8.888	61.888	-5.888	
59	VAL	CB	-8.888	61.888	-5.888	
59	VAL	CG1	-8.888	61.888	-5.888	
59	VAL	CG2	-8.888	61.888	-5.888	
60	PRO	N	-8.888	61.888	-5.888	
60	PRO	C	-8.888	61.888	-5.888	
60	PRO	CB	-8.888	61.888	-5.888	
60	PRO	CG	-8.888	61.888	-5.888	
60	PRO	CD	-8.888	61.888	-5.888	
61	ASP	OD2	-8.888	61.888	-5.888	
61	ASP	CG	-8.888	61.888	-5.888	
61	ASP	CA	-8.888	61.888	-5.888	
61	ASP	O	-8.888	61.888	-5.888	
62	LEU	CA	-8.888	61.888	-5.888	
62	LEU	O	-8.888	61.888	-5.888	
62	LEU	CG	-8.888	61.888	-5.888	
62	LEU	CD2	-8.888	61.888	-5.888	
63	LYS	CA	-8.888	61.888	-5.888	
63	LYS	O	-8.888	61.888	-5.888	
63	LYS	CG	-8.888	61.888	-5.888	
63	LYS	CE	-8.888	61.888	-5.888	
64	VAL	N	-8.888	61.888	-5.888	
64	VAL	C	-8.888	61.888	-5.888	
64	VAL	CB	-8.888	61.888	-5.888	
64	VAL	CG2	-8.888	61.888	-5.888	
65	ALA	CA	-8.888	61.888	-5.888	
65	ALA	O	-8.888	61.888	-5.888	
66	GLY	N	-8.888	61.888	-5.888	
66	GLY	C	-8.888	61.888	-5.888	
67	GLY	N	-8.888	61.888	-5.888	
67	GLY	C	-8.888	61.888	-5.888	
68	ALA	N	-8.888	61.888	-5.888	
68	ALA	C	-8.888	61.888	-5.888	
68	ALA	CB	-8.888	61.888	-5.888	
69	SER	CA	-8.888	61.888	-5.888	
69	SER	O	-8.888	61.888	-5.888	
69	SER	OG	-8.888	61.888	-5.888	
70	MET	CA	-8.888	61.888	-5.888	
70	MET	O	-8.888	61.888	-5.888	
70	MET	CG	-8.888	61.888	-5.888	
70	MET	CE	-8.888	61.888	-5.888	
71	VAL	CA	-8.888	61.888	-5.888	
71	VAL	O	-8.888	61.888	-5.888	
71	VAL	CG1	-8.888	61.888	-5.888	
72	PRO	N	-8.888	61.888	-5.888	
72	PRO	C	-8.888	61.888	-5.888	
72	PRO	CB	-8.888	61.888	-5.888	
72	PRO	CD	-8.888	61.888	-5.888	
73	SER	CA	-8.888	61.888	-5.888	
73	SER	O	-8.888	61.888	-5.888	
73	SER	OG	-8.888	61.888	-5.888	
74	GLU	CA	-8.888	61.888	-5.888	
74	GLU	O	-8.888	61.888	-5.888	
74	GLU	CG	-8.888	61.888	-5.888	
74	GLU	OE1	-8.888	61.888	-5.888	

54	GLN NE2	-3.988	55.777	8.271	
55	THR CA	-8.633	58.121	-5.441	
55	THR B	-9.433	57.919	-7.810	
55	THR OG1	-9.885	60.310	-5.418	
56	ASN N	-7.482	58.603	-4.877	
56	ASN OD1	-5.875	58.967	-10.337	
56	ASN CB	-5.898	59.894	-8.208	
56	ASN C	-4.812	57.894	-8.305	
57	PRO N	-6.362	56.261	-9.298	
57	PRO CD	-7.384	54.633	-10.272	
57	PRO CA	-5.679	56.961	-9.332	
57	PRO B	-3.509	54.228	-9.945	
58	PHE CA	-2.747	54.577	-11.222	
58	PHE B	-0.635	57.697	-10.680	
58	PHE CG	-3.983	58.968	-13.357	
58	PHE CD2	-5.211	57.838	-13.459	
58	PHE CE2	-6.194	57.895	-14.274	
59	GLN N	-2.844	57.119	-8.998	
59	GLN C	-0.807	56.403	-7.868	
59	GLN CB	-1.862	58.648	-7.889	
59	GLN CG	-1.798	60.157	-5.159	
59	GLN NE2	-2.959	59.685	-6.762	
60	ASP CA	0.851	54.792	-6.384	
60	ASP B	2.827	55.550	-5.231	
60	ASP CG	2.677	52.338	-6.388	
60	ASP OD2	2.915	51.841	-7.838	
61	ASN ND2	-1.364	57.747	-2.347	
61	ASN CG	-0.648	57.678	-2.399	
61	ASN CA	1.557	55.734	-2.789	
61	ASN B	2.933	54.862	-0.982	
62	ASN CA	2.877	52.348	-1.709	
62	ASN B	4.951	51.313	-1.778	
62	ASN CG	2.371	50.183	-0.897	
62	ASN OD2	2.622	50.258	0.601	
63	SER CA	5.189	51.696	-0.709	
63	SER B	5.593	49.790	-0.269	
63	SER OG	6.871	50.498	-3.418	
64	HIS CA	3.994	48.859	-0.935	
64	HIS B	3.861	46.974	-7.108	
64	HIS CG	3.144	46.821	-3.726	
64	HIS CD2	4.854	45.194	-3.135	
64	HIS NE2	3.554	43.920	-3.368	
65	GLY CA	1.552	48.264	-7.838	
65	GLY O	2.238	48.878	-10.134	
66	THR CA	4.864	50.117	-9.954	
66	THR B	5.333	48.789	-11.661	
66	THR OG1	3.637	52.435	-9.606	
67	HIS N	5.685	50.663	-9.274	
67	HIS C	6.891	46.141	-10.143	
67	HIS CB	7.308	47.871	-8.064	
67	HIS ND1	8.598	44.957	-8.278	
67	HIS CE1	9.857	44.491	-8.299	
68	VAL N	4.892	45.749	-9.731	
68	VAL C	8.856	44.868	-11.748	
68	VAL CB	2.939	44.252	-9.384	
68	VAL CG2	3.319	43.785	-8.888	
69	ALA CA	8.837	44.468	-13.429	
69	ALA B	6.828	43.913	-15.565	
70	GLY N	5.348	46.782	-13.916	
70	GLY C	9.846	43.378	-15.821	
71	THR N	6.828	44.431	-14.138	
71	THR C	8.224	42.586	-18.543	
71	THR CB	7.119	42.878	-13.191	

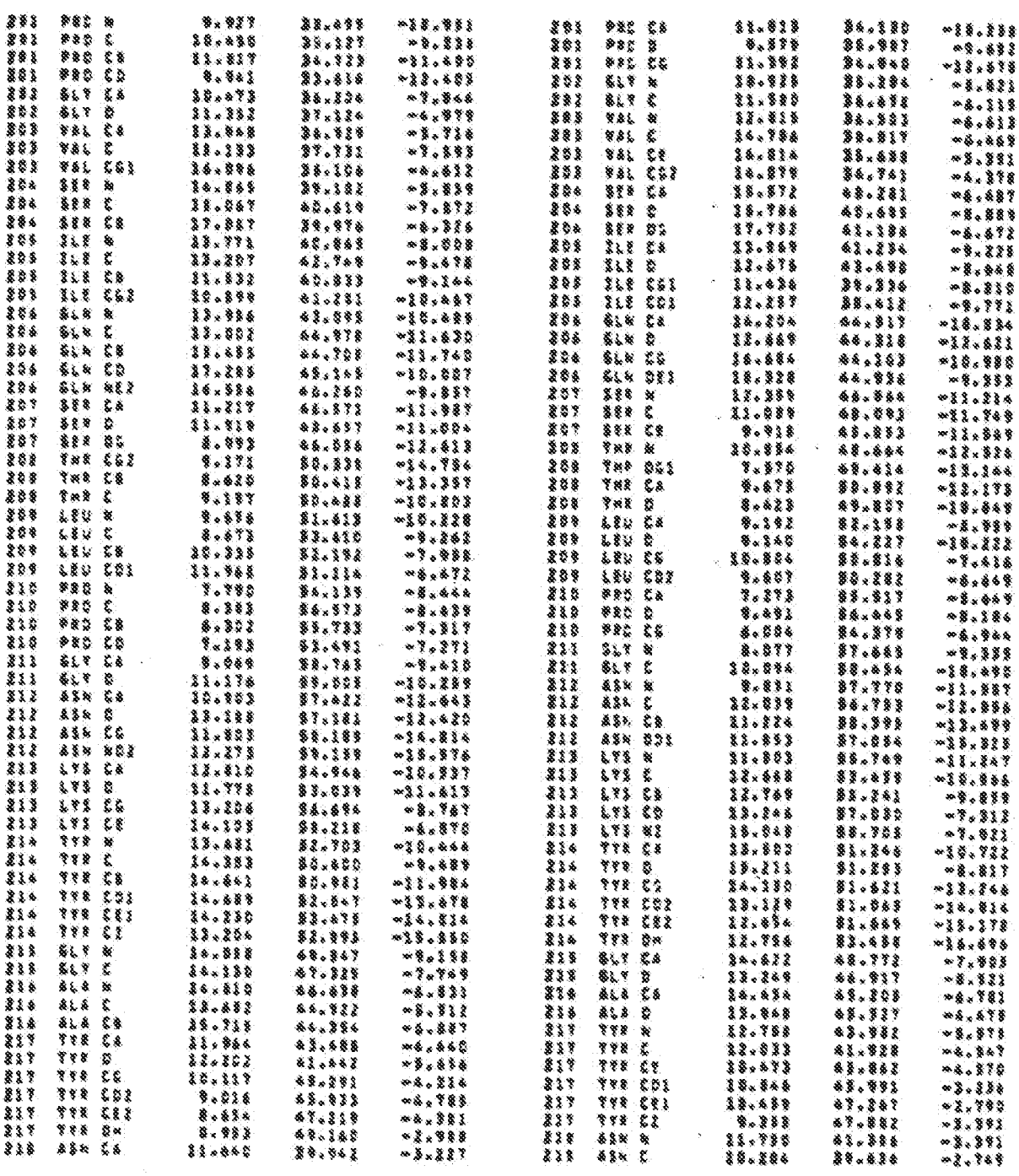
51	THR N	-8.571	58.251	-6.249	
51	THR C	-8.784	58.139	-8.779	
51	THR CB	-10.386	59.288	-5.383	
51	THR CG2	-11.432	59.143	-6.817	
54	ASN ND2	-6.938	61.179	-9.881	
54	ASN CG	-6.273	59.925	-9.555	
54	ASN CA	-6.742	58.425	-8.288	
54	ASN O	-5.184	56.866	-7.678	
57	PRO CG	-7.123	55.257	-11.177	
57	PRO CB	-6.044	54.178	-10.235	
57	PRO C	-4.381	55.982	-9.944	
58	PHE N	-3.998	56.262	-10.491	
58	PHE C	-1.712	57.128	-10.253	
58	PHE CB	-2.943	57.582	-12.423	
58	PHE CD1	-3.756	55.788	-14.059	
58	PHE CE1	-4.722	55.255	-14.928	
58	PHE CZ	-5.949	55.939	-15.051	
59	GLN CA	-1.172	57.583	-7.934	
59	GLN O	-1.639	56.883	-6.115	
59	GLN CG	-0.942	59.261	-6.834	
59	GLN OE1	-1.084	61.288	-6.836	
60	ASP N	0.418	55.895	-7.211	
60	ASP C	1.631	55.267	-5.899	
60	ASP CB	1.394	53.744	-7.188	
60	ASP OD1	1.744	52.837	-5.190	
61	ASN N	0.959	55.263	-3.950	
61	ASN OD1	0.666	58.546	-2.875	
61	ASN CB	0.531	56.481	-1.784	
61	ASN C	2.291	54.632	-1.948	
62	ASN N	2.219	53.436	-2.448	
62	ASN C	4.324	51.893	-2.479	
62	ASN CB	1.783	51.319	-1.421	
62	ASN OD1	2.633	49.877	-1.343	
63	SER N	4.352	52.106	-3.741	
63	SER C	5.871	50.254	-5.289	
63	SER CB	6.523	51.958	-4.812	
64	HIS N	4.282	49.475	-4.839	
64	HIS C	3.366	47.759	-6.241	
64	HIS CB	3.184	47.581	-3.747	
64	HIS ND1	2.187	45.247	-4.241	
64	HIS CE1	2.616	43.068	-4.854	
65	GLY N	2.287	48.428	-6.587	
65	GLY C	2.391	48.636	-9.837	
66	THR N	3.233	49.659	-8.832	
66	THR C	5.889	49.809	-10.291	
66	THR CB	4.744	51.513	-9.687	
66	THR CG2	5.536	52.878	-10.849	
67	HIS CA	6.783	47.341	-9.458	
67	HIS O	6.449	45.638	-11.350	
67	HIS CG	8.595	46.275	-8.168	
67	HIS CD2	9.904	46.678	-8.976	
67	HIS NE2	10.678	45.514	-8.108	
68	VAL CA	6.142	44.887	-10.286	
68	VAL O	4.114	43.962	-12.335	
68	VAL CG1	3.988	43.248	-10.028	
69	ALA N	5.373	46.849	-12.113	
69	ALA C	8.193	46.398	-14.411	
69	ALA CB	2.332	47.851	-13.386	
70	GLY CA	6.595	46.805	-16.878	
70	GLY O	7.884	45.154	-16.119	
71	THR CA	7.177	43.819	-14.644	
71	THR O	6.682	41.828	-16.495	
71	THR OG1	8.191	42.592	-12.398	

71	TYR	CG2	7.274	48.983	-13.996	72	VAL	N	6.938	42.887	-15.427
72	VAL	CA	3.976	42.491	-16.696	72	VAL	C	6.312	43.884	-17.831
72	VAL	N	4.341	42.388	-18.888	72	VAL	CB	2.916	42.867	-16.885
72	VAL	CG1	1.313	42.498	-17.178	72	VAL	CG2	2.142	42.327	-14.723
73	ALA	N	6.584	44.417	-17.888	73	ALA	CA	6.587	43.891	-19.147
73	ALA	C	8.433	46.333	-19.355	73	ALA	O	9.862	47.188	-20.216
73	ALA	CB	8.187	45.441	-19.433	74	ALA	N	6.544	46.429	-18.435
74	ALA	CA	7.478	47.581	-18.955	74	ALA	C	7.740	47.448	-20.342
74	ALA	N	7.959	48.640	-21.884	74	ALA	CB	8.653	47.444	-17.925
75	LEU	N	7.658	48.784	-21.839	75	LEU	CA	7.812	48.968	-22.458
75	LEU	C	9.192	48.568	-22.966	75	LEU	O	10.162	48.758	-23.253
75	LEU	CB	7.548	50.471	-22.809	75	LEU	CG	6.123	50.913	-22.379
75	LEU	CD1	6.079	52.436	-22.380	75	LEU	CD2	5.896	50.442	-23.685
76	ASN	N	9.167	48.183	-24.169	76	ASN	ND2	12.385	46.452	-26.384
76	ASN	OD1	10.950	45.860	-27.878	76	ASN	CG	11.195	46.276	-26.802
76	ASN	CB	10.810	46.851	-25.988	76	ASN	CA	10.359	47.738	-24.938
76	ASN	C	10.783	49.848	-25.643	76	ASN	O	10.157	49.479	-26.419
77	ASN	N	11.804	49.664	-25.871	77	ASN	CA	12.220	50.957	-25.681
77	ASN	C	13.787	51.829	-25.348	77	ASN	O	14.364	49.979	-25.313
77	ASN	CB	11.335	52.876	-25.117	77	ASN	CG	11.258	52.827	-23.616
77	ASN	OD1	12.032	51.346	-22.917	77	ASN	ND2	10.294	53.741	-23.825
78	SER	N	14.125	52.287	-25.144	78	SER	CA	15.513	52.634	-24.906
78	SER	C	15.818	52.742	-23.436	78	SER	O	16.982	53.871	-23.144
78	SER	CB	15.905	53.941	-25.587	78	SER	OG	15.926	53.878	-26.999
79	ILE	N	14.858	52.565	-22.529	79	ILE	CA	15.195	52.786	-21.120
79	ILE	C	14.617	51.883	-20.238	79	ILE	O	13.843	53.841	-20.679
79	ILE	CB	14.471	54.174	-20.697	79	ILE	CG1	12.945	54.832	-20.814
79	ILE	CG2	14.997	55.328	-21.612	79	ILE	CD1	12.135	55.176	-20.155
80	GLY	N	14.995	51.768	-18.981	80	GLY	CA	14.476	50.948	-17.913
80	GLY	C	14.612	49.448	-18.219	80	GLY	O	15.719	48.994	-18.346
81	VAL	N	13.513	48.766	-17.980	81	VAL	CA	13.411	47.286	-18.041
81	VAL	C	12.511	46.919	-19.217	81	VAL	O	12.260	47.739	-20.117
81	VAL	CB	13.001	46.755	-16.677	81	VAL	CG1	14.030	47.884	-15.573
81	VAL	CG2	11.638	47.281	-16.231	82	LEU	N	12.126	45.645	-19.216
82	LEU	CA	11.312	45.828	-20.256	82	LEU	C	10.390	44.828	-19.518
82	LEU	O	10.858	43.356	-18.650	82	LEU	CB	12.284	44.219	-21.229
82	LEU	CG	11.439	43.568	-22.366	82	LEU	CD1	10.796	44.657	-23.223
82	LEU	CD2	12.359	42.675	-23.192	83	GLY	N	9.131	44.180	-19.816
83	GLY	CA	8.833	43.321	-19.114	83	GLY	C	8.827	42.811	-20.925
83	GLY	O	8.546	41.822	-21.026	84	VAL	N	7.272	41.112	-19.283
84	VAL	CA	6.973	39.807	-19.888	84	VAL	C	6.144	40.839	-21.148
84	VAL	O	6.424	39.472	-22.196	84	VAL	CB	6.254	38.920	-18.861
84	VAL	CG1	5.680	37.677	-19.557	84	VAL	CG2	7.590	38.587	-17.705
85	ALA	N	5.158	40.926	-21.024	85	ALA	CA	4.217	41.196	-22.158
85	ALA	C	4.213	42.683	-22.396	85	ALA	O	3.240	43.481	-22.838
85	ALA	CB	2.846	40.663	-21.768	86	PRO	N	5.240	43.186	-23.858
86	PRO	CA	5.413	44.635	-23.285	86	PRO	C	6.321	45.371	-23.947
86	PRO	O	6.291	44.825	-25.849	86	PRO	CB	6.822	44.784	-23.813
86	PRO	CG	7.838	43.466	-24.544	86	PRO	CD	6.377	42.468	-23.634
87	SER	N	3.548	46.676	-24.769	87	SER	CA	2.489	45.324	-25.529
87	SER	C	1.103	45.132	-24.897	87	SER	O	0.362	45.513	-25.619
87	SER	CB	2.401	44.777	-26.827	87	SER	OG	3.591	45.143	-27.583
88	ALA	N	1.817	44.584	-23.762	88	ALA	CB	-0.163	43.818	-21.828
88	ALA	CA	-0.273	44.353	-23.684	88	ALA	C	-0.898	45.717	-22.898
88	ALA	O	-0.174	46.717	-22.435	89	SER	N	-2.219	45.871	-22.678
89	SER	OG	-4.144	47.102	-24.280	89	SER	CB	-4.343	46.983	-22.898
89	SER	CA	-2.801	46.867	-22.227	89	SER	C	-3.136	46.788	-20.727
89	SER	O	-3.793	45.844	-20.209	90	LEU	N	-2.648	47.658	-20.837
90	LEU	CA	-2.378	47.447	-18.983	90	LEU	C	-3.483	48.438	-17.864
90	LEU	O	-3.582	48.404	-18.215	90	LEU	CB	-0.951	48.273	-18.424
90	LEU	CG	-0.233	47.851	-17.174	90	LEU	CD1	-0.826	48.341	-17.219
90	LEU	CD2	1.160	48.524	-17.047	91	TYR	N	-4.284	47.946	-16.938
91	TYR	CA	-5.258	48.478	-16.137	91	TYR	C	-4.873	48.750	-34.485

91	TYR	CB	-6.686	48.893	-16.314
91	TYR	CD1	-6.596	47.415	-18.755
91	TYR	CE1	-6.985	47.572	-20.098
91	TYR	CZ	-7.796	48.582	-20.443
92	ALA	N	-4.895	49.958	-14.184
92	ALA	C	-5.823	50.833	-11.963
92	ALA	CB	-3.997	51.621	-12.488
93	VAL	CA	-7.183	48.856	-10.325
93	VAL	O	-6.181	47.993	-8.372
93	VAL	CG1	-9.213	47.488	-9.725
94	LYS	N	-6.987	50.217	-8.327
94	LYS	C	-7.331	49.985	-5.894
94	LYS	CB	-8.851	51.976	-6.818
94	LYS	CD	-9.868	53.785	-5.582
94	LYS	NZ	-8.735	55.544	-4.387
95	VAL	CA	-7.646	48.457	-3.920
95	VAL	O	-7.425	48.136	-1.581
95	VAL	CG1	-8.868	46.852	-5.619
96	LEU	N	-5.474	48.974	-2.684
96	LEU	C	-4.331	50.559	-1.321
96	LEU	CB	-3.589	48.241	-1.573
96	LEU	CD1	-2.287	46.184	-2.163
97	GLY	N	-4.324	50.975	-0.886
97	GLY	C	-2.363	52.437	0.385
98	ALA	N	-1.954	53.648	0.758
98	ALA	CA	-0.563	54.868	0.945
98	ALA	O	1.393	52.921	1.663
99	ASP	OD2	-2.631	51.842	4.191
99	ASP	CG	-2.883	51.131	3.848
99	ASP	CA	0.181	53.418	3.855
99	ASP	O	0.735	49.313	4.829
100	GLY	CA	-0.343	48.521	1.415
100	GLY	O	-1.649	46.512	1.479
101	SER	CA	-3.542	47.388	3.315
101	SER	O	-4.758	48.972	1.987
101	SER	OG	-4.411	48.434	5.289
102	GLY	CA	-7.877	47.422	1.896
102	GLY	O	-7.888	45.431	3.838
103	GLN	CA	-10.535	46.297	3.820
103	GLN	C	-10.779	45.482	0.817
103	GLN	CG	-11.348	48.885	4.586
103	GLN	OE1	-12.159	48.816	5.902
104	TYR	N	-11.811	44.141	2.451
104	TYR	C	-13.831	43.690	0.473
104	TYR	CB	-12.697	41.846	2.109
104	TYR	CD1	-11.819	39.789	3.377
104	TYR	CE1	-10.805	38.885	3.787
104	TYR	CZ	-9.564	39.822	3.881
105	SER	N	-13.589	44.572	0.983
105	SER	C	-14.172	45.820	-1.159
105	SER	CB	-15.880	44.121	0.601
106	TRP	N	-13.879	46.425	-2.834
106	TRP	C	-13.895	46.436	-3.833
106	TRP	CB	-11.321	48.254	-3.355
106	TRP	CD1	-12.842	49.526	0.264
106	TRP	NE1	-12.691	50.358	1.368
106	TRP	CE3	-9.375	49.852	0.576
106	TRP	CZ3	-8.568	50.563	1.525
107	ILE	N	-11.339	45.338	-2.481
107	ILE	C	-11.955	43.584	-4.198
107	ILE	CB	-9.944	43.183	-2.523
107	ILE	CG2	-9.632	41.938	-3.381
107	ILE	N	-12.894	43.292	-3.977

91	TYR	O	-6.496	47.749	-14.823
91	TYR	CG	-7.894	48.237	-17.761
91	TYR	CD2	-7.971	49.275	-18.149
91	TYR	CE2	-8.315	49.421	-19.492
91	TYR	OH	-8.182	48.752	-21.764
92	ALA	CA	-4.949	50.199	-12.787
92	ALA	O	-6.723	52.878	-12.850
93	VAL	N	-5.959	48.993	-11.329
93	VAL	C	-4.708	50.814	-8.899
93	VAL	CB	-7.957	47.555	-10.431
93	VAL	CG2	-8.195	47.378	-12.872
94	LYS	CA	-6.378	50.464	-6.999
94	LYS	O	-6.458	50.480	-5.783
94	LYS	CG	-5.896	52.320	-5.467
94	LYS	CE	-4.389	54.208	-4.189
95	VAL	N	-4.909	49.071	-5.026
95	VAL	C	-6.919	48.499	-2.568
95	VAL	CB	-8.184	47.838	-4.319
95	VAL	CG2	-6.980	46.180	-4.332
96	LEU	CA	-4.782	49.183	-1.486
96	LEU	O	-3.942	51.121	-2.336
96	LEU	CG	-3.593	46.799	-2.072
96	LEU	CD2	-4.459	46.882	-1.845
97	GLY	CA	-3.898	52.387	0.287
97	GLY	O	-1.619	51.463	0.165
98	ALA	CB	-0.428	55.478	1.510
98	ALA	C	0.188	53.118	1.917
99	ASP	N	-0.504	52.573	2.912
99	ASP	OD1	-2.738	50.982	4.883
99	ASP	CB	-0.448	51.683	5.175
99	ASP	C	0.146	50.185	3.328
100	GLY	N	-0.424	49.883	2.168
100	GLY	C	-1.528	47.851	2.802
101	SER	N	-2.542	48.128	2.988
101	SER	C	-4.750	47.894	2.532
101	SER	CB	-3.716	47.447	4.817
102	GLY	N	-5.821	47.892	2.577
102	GLY	C	-8.166	46.536	2.528
103	GLN	N	-9.377	47.858	2.498
103	GLN	C	-10.963	45.232	2.022
103	GLN	CB	-11.471	47.307	3.274
103	GLN	CD	-12.388	49.184	4.915
103	GLN	NE2	-13.419	49.157	4.312
104	TYR	CA	-12.868	43.126	1.508
104	TYR	O	-12.939	43.278	-0.887
104	TYR	CG	-11.629	40.829	2.472
104	TYR	CD2	-10.379	40.959	1.868
104	TYR	CE2	-9.352	40.057	2.271
104	TYR	OH	-8.481	38.181	3.324
105	SER	CA	-14.877	45.166	-0.834
105	SER	O	-14.759	45.935	-2.258
105	SER	OG	-15.289	47.839	1.450
106	TRP	CA	-12.421	47.391	-1.948
106	TRP	O	-12.821	46.448	-4.245
106	TRP	CG	-11.845	49.111	-0.206
106	TRP	CD2	-10.458	49.812	0.981
106	TRP	CE2	-11.359	50.973	1.581
106	TRP	CZ2	-10.671	51.318	2.580
106	TRP	CH2	-9.293	51.291	2.455
107	ILE	CA	-10.785	44.258	-3.325
107	ILE	O	-11.895	43.474	-5.398
107	ILE	CG1	-8.434	43.784	-1.924
107	ILE	CD1	-8.243	42.998	-0.427

```
108  ILE CA    -14.116   42.722   -4.321     108  ILE C     -14.839   43.694   -5.386
108  ILE O     -14.896   43.328   -6.552     108  ILE CB    -15.246   42.265   -3.328
108  ILE CG1   -14.726   41.077   -2.482     108  ILE CG2   -16.568   42.824   -4.893
108  ILE CD1   -15.432   40.845   -1.131     109  ASN N     -14.781   44.958   -4.981
109  ASN CA    -15.204   46.019   -5.916     109  ASN C     -14.232   46.867   -7.886
109  ASN O     -14.660   48.272   -8.235     109  ASN CB    -15.280   47.359   -5.207
109  ASN CG    -16.578   47.600   -4.353     109  ASN OD1   -17.495   46.695   -4.646
109  ASN ND2   -16.433   48.447   -3.442     110  GLY N     -12.951   45.908   -6.774
110  GLY CA    -11.992   45.917   -7.865     110  GLY C     -12.108   44.712   -8.812
110  GLY O     -11.929   44.929  -10.034     111  ILE N     -12.379   43.539   -8.246
111  ILE CA    -12.603   42.336   -9.099     111  ILE C     -13.859   42.560   -9.942
111  ILE O     -13.921   42.384  -11.148     111  ILE CB    -12.734   40.968   -8.364
111  ILE CG1   -11.421   40.501   -7.655     111  ILE CG2   -13.122   39.791   -9.367
111  ILE CD1   -11.588   39.786   -6.336     112  GLU N     -14.893   43.875   -9.280
112  GLU CA    -16.118   43.376  -10.846     112  GLU C     -15.872   44.367  -11.171
112  GLU O     -16.447   44.130  -12.246     112  GLU CB    -17.229   43.899   -9.141
112  GLU CG    -17.847   42.917   -8.135     112  GLU CD    -18.724   41.826   -8.685
112  GLU OE1   -19.841   40.866   -8.016     112  GLU OE2   -19.123   41.928   -9.866
113  TRP N     -15.096   45.403  -10.971     113  TRP CA    -14.756   46.408  -12.000
113  TRP C     -14.076   45.663  -13.140     113  TRP O     -14.319   45.932  -14.332
113  TRP CB    -13.882   47.553  -11.434     113  TRP CG    -13.486   48.556  -12.481
113  TRP CD1   -14.148   49.736  -12.681     113  TRP CD2   -12.441   48.551  -13.463
113  TRP NE1   -13.597   50.443  -13.723     113  TRP CE2   -12.565   49.761  -14.215
113  TRP CE3   -11.451   47.645  -13.809     113  TRP CZ2   -11.696   50.845  -15.276
113  TRP CZ3   -10.610   47.899  -14.879     113  TRP CH2   -10.752   49.074  -15.603
114  ALA N     -13.089   44.801  -12.832     114  ALA CA    -12.333   44.865  -13.874
114  ALA C     -13.199   43.179  -14.752     114  ALA O     -12.943   43.874  -15.978
114  ALA CB    -11.299   43.182  -13.160     115  ILE N     -14.174   42.549  -14.119
115  ILE CA    -15.870   41.648  -14.897     115  ILE C     -15.928   42.485  -15.856
115  ILE O     -16.077   42.225  -17.070     115  ILE CB    -16.880   40.840  -13.922
115  ILE CG1   -15.218   39.836  -13.843     115  ILE CG2   -17.151   40.168  -14.755
115  ILE CD1   -16.004   39.411  -11.743     116  ALA N     -16.534   43.327  -15.267
116  ALA CA    -17.390   44.460  -16.850     116  ALA C     -16.786   45.069  -17.278
116  ALA O     -17.323   45.255  -18.343     116  ALA CB    -18.011   45.310  -15.181
117  ASN N     -15.423   45.390  -17.122     117  ASN CA    -14.353   45.987  -18.139
117  ASN C     -13.827   44.974  -19.834     117  ASN O     -12.997   45.436  -19.820
117  ASN CB    -13.815   46.958  -17.426     117  ASN CG    -14.400   48.177  -16.939
117  ASN OD1   -14.385   49.882  -17.773     117  ASN ND2   -14.931   48.269  -15.736
118  ASN N     -14.223   43.725  -18.967     118  ASN CA    -13.740   42.642  -19.832
118  ASN C     -12.240   42.444  -19.863     118  ASN O     -11.817   42.309  -20.932
118  ASN CB    -14.247   42.863  -21.279     118  ASN CG    -15.737   43.080  -21.395
118  ASN OD1   -16.510   42.321  -20.759     118  ASN ND2   -16.136   44.096  -22.133
119  MET N     -11.686   42.500  -18.675     119  MET CA    -10.232   41.222  -18.478
119  MET C     -10.825   40.734  -18.928     119  MET O     -10.888   39.838  -18.759
119  MET CB     -9.819   42.461  -17.053     119  MET CG     -9.880   43.883  -16.382
119  MET SD     -8.788   44.963  -17.526     119  MET CE     -9.982   46.061  -18.263
120  ASP N      -8.904   40.437  -19.584     120. ASP CA     -8.480   39.118  -20.830
120  ASP C      -7.822   38.390  -18.856     120  ASP O      -8.838   37.189  -18.690
120  ASP CB     -7.555   39.156  -21.236     120  ASP CG     -8.237   39.730  -22.434
120  ASP OD1    -7.881   40.706  -23.886     120  ASP OD2    -9.327   39.135  -22.739
121  VAL N      -7.071   39.117  -18.115     121  VAL CA     -6.224   38.681  -16.974
121  VAL C      -6.296   39.534  -15.786     121  VAL O      -6.284   40.788  -15.989
121  VAL CB     -4.755   38.507  -17.696     121  VAL CG1    -3.758   38.176  -16.627
121  VAL CG2    -4.787   37.916  -18.846     122  ILE N      -6.318   38.978  -14.598
122  ILE CA     -6.268   39.799  -13.397     122  ILE C      -5.828   39.262  -12.627
122  ILE O      -6.829   38.012  -12.469     122  ILE CB     -7.476   39.606  -12.466
122  ILE CG1    -8.686   40.392  -13.863     122  ILE CG2    -7.221   39.883  -10.956
122  ILE CD1    -9.976   39.788  -12.383     123  ASN N      -4.263   40.222  -12.119
123  ASN CA     -3.145   39.854  -11.232     123  ASN C      -3.582   40.606   -9.861
123  ASN O      -3.788   41.831   -9.833     123  ASN CB     -1.828   40.478  -11.497
123  ASN CG     -0.692   40.868  -10.777     123  ASN OD1    -0.863   38.990  -11.818
123  ASN ND2    -0.366   40.747   -9.728     124  MET N      -3.458   39.604   -8.332
124  MET CA     -3.658   39.973   -7.438     124  MET C      -2.623   39.683   -6.814
```



268	ASN	CA	1.353	27.973	-26.497	269	ASN	C	0.838	28.994	-28.681
269	ASN	O	0.943	29.782	-25.943	269	ASN	CB	0.403	26.813	-27.895
269	ASN	CG	3.161	26.826	-25.211	269	ASN	OD1	0.593	27.826	-25.181
269	ASN	ND2	25.831	26.796	-25.973	270	VAL	N	0.888	28.868	-29.784
270	VAL	CA	2.843	31.418	-29.819	270	VAL	C	0.888	28.887	-31.814
270	VAL	O	3.897	29.949	-29.872	270	VAL	CB	0.898	31.918	-29.432
270	VAL	CG1	6.868	32.797	-31.899	270	VAL	CG2	0.438	32.362	-28.282
271	GLN	N	3.328	29.781	-32.892	271	GLN	CA	7.682	29.378	-34.764
271	GLN	C	6.889	27.934	-35.831	271	GLN	O	8.313	27.288	-36.891
271	GLN	CB	5.184	29.218	-34.864	271	GLN	CG	9.486	28.618	-36.336
271	GLN	CD	28.901	28.313	-36.892	271	GLN	OE1	31.869	28.378	-37.718
271	GLN	NE2	31.782	28.833	-38.818	272	ALA	N	6.877	26.909	-34.892
272	ALA	CA	6.854	25.712	-34.142	272	ALA	C	6.781	23.958	-34.146
272	ALA	O	5.878	23.503	-34.851	272	ALA	CB	6.963	26.763	-32.172
273	ALA	N	8.267	24.861	-33.285	273	ALA	CA	8.748	26.931	-32.859
273	ALA	C	9.861	22.323	-34.020	273	ALA	O	8.889	23.219	-34.888
273	ALA	CB	9.736	24.773	-31.885	274	ALA	N	9.765	24.464	-34.944
274	ALA	CA	8.953	25.391	-30.810	274	ALA	CA	8.189	29.164	-34.847
274	ALA	C	8.738	25.367	-29.290	274	ALA	O	8.888	28.949	-31.833
275	GLN	N	8.838	27.154	-28.714	275	GLN	CA	8.848	28.389	-28.837
275	GLN	C	8.147	27.241	-27.777	275	GLN	O	9.988	27.827	-26.918
275	GLN	CB	9.193	27.361	-28.695	275	GLN	CD	8.836	28.914	-28.128
275	GLN	CG	8.921	24.664	-27.447	275	GLN	OE1	-0.823	29.986	-27.832
275	GLN	NE2	-1.378	28.879	-28.729	275	GLN	HE2	-1.378	28.818	-26.838

The above structural studies together with the kinetic data presented herein and elsewhere (Philipp, M., et al. (1983) Mol. Cell. Biochem. 51, 5-32; Svendsen, I.B. (1976) Carlsberg Res. Comm. 41, 237-291; Markland, S.F. Id; Stauffe, D.C., et al. (1965) J. Biol. Chem. 244, 5333-5338) indicate that the subsites in the binding cleft of subtilisin are capable of interacting with substrate amino acid residues from P-4 to P-2'.

The most extensively studied of the above residues are Gly166, Gly169 and Ala152. These amino acids were identified as residues within the S-1 subsite. As seen in Fig. 3, which is a stereoview of the S-1 subsite, Gly166 and Gly169 occupy positions at the bottom of the S-1 subsite, whereas Ala152 occupies a position near the top of S-1, close to the catalytic Ser221.

All 19 amino acid substitutions of Gly166 and Gly169 have been made. As will be indicated in the examples which follow, the preferred replacement amino acids for Gly166 and/or Gly169 will depend on the specific amino acid occupying the P-1 position of a given substrate.

The only substitutions of Ala152 presently made and analyzed comprise the replacement of Ala152 with Gly and Ser. The results of these substitutions on P-1 specificity will be presented in the examples.

In addition to those residues specifically associated with specificity for the P-1 substrate amino acid, Tyr104 has been identified as being involved with P-4 specificity. Substitutions at Phe189 and Tyr217,

however, are expected to respectively effect P-2' and P-1' specificity.

The catalytic activity of subtilisin has also been modified by single amino acid substitutions at Asn155. The catalytic triad of subtilisin is shown in Fig. 4. As can be seen, Ser221, His64 and Asp32 are positioned to facilitate nucleophilic attach by the serine hydoxylate on the carbonyl of the scissile peptide bond. Crystallographic studies of subtilisin (Robertus, et al. (1972) Biochem. 11, 4293-4303; Matthews, et al. (1975) J. Biol. Chem. 250, 7120-7126; Poulos, et al. (1976) J. Biol. Chem. 250, 1097-1103) show that two hydrogen bonds are formed with the oxyanion of the substrate transition state. One hydrogen bond donor is from the catalytic serine-221 main-chain amide while the other is from one of the NE2 protons of the asparagine-155 side chain. See Fig. 4.

Asn155 was substituted with Ala, Asp, His, Glu and Thr. These substitutions were made to investigate the the stabilization of the charged tetrahedral intermediate of the transition state complex by the potential hydrogen bond between the side chain of Asn155 and the oxyanion of the intermediate. These particular substitutions caused large decreases in substrate turnover, Kcat (200 to 4,000 fold), marginal decreases in substrate binding Km (up to 7 fold), and a loss in transition state stabilization energy of 2.2 to 4.7 Kcal/mol. The retention of Km and the drop in Kcat will make these mutant enzymes useful as binding proteins for specific peptide sequences, the nature of which will be determined by the specificity of the precursor protease.

Various other amino acid residues have been identified which affect alkaline stability. In some cases, mutants having altered alkaline stability also have altered thermal stability.

In B amyloliquefaciens subtilisin residues Asp36, Ile107, Lys170, Ser204 and Lys213 have been identified as residues which upon substitution with a different amino acid alter the alkaline stability of the mutated enzyme as compared to the precursor enzyme. The substitution of Asp36 with Ala and the substitution of Lys170 with Glu each resulted in a mutant enzyme having a lower alkaline stability as compared to the wild type subtilisin. When Ile107 was substituted with Val, Ser204 substituted with Cys, Arg or Leu or Lys213 substituted with Arg, the mutant subtilisin had a greater alkaline stability as compared to the wild type subtilisin. However, the mutant Ser204P demonstrated a decrease in alkaline stability.

In addition, other residues, identified as being associated with the modification of other properties of subtilisin, also affect alkaline stability. These residues include Ser24, Met50, Glu156, Gly166, Gly169 and Tyr217. Specifically the following particular substitutions result in an increased alkaline stability: Ser24C, Met50F, Gly156Q or S, Gly166A, H, K, N or Q, Gly169S or A, and Tyr217F, K, R or L. The mutant Met50V, on the other hand, results in a decrease in the alkaline stability of the mutant subtilisin as compared to wild type subtilisin.

Other residues involved in alkaline stability based on the alkaline stability screen include Asp197 and Met222. Particular mutants include Asp197(R or A) and Met 222 (all other amino acids).

Various other residues have been identified as being involved in thermal stability as determined by the thermal stability screen herein. These residues include the above identified residues which effect alkaline stability and Met199 and Tyr21. These latter two residues are also believed to be important for alkaline stability. Mutants at these residues include I199 and F21.

The amino acid sequence of B. amyloliquefaciens substilisin has also been modified by substituting two or more amino acids of the wild-type sequence. Six categories of multiply substituted mutant subtilisin have been identified. The first two categories comprise thermally and oxidatively stable mutants. The next three other categories comprise mutants which combine the useful properties of any of several single mutations of B. amyloliquefaciens subtilisin. The last category comprises mutants which have modified alkaline and/or thermal stability.

The first category comprises double mutants in which two cysteine residues have been substituted at various amino acid residue positions within the subtilisin molecule. Formation of disulfide bridges between the two substituted cysteine residues results in mutant subtilisins with altered thermal stability and catalytic activity. These mutants include A21/C22/C87 and C24/C87 which will be described in more detail in Example 11.

The second category of multiple subtilisin mutants comprises mutants which are stable in the presence of various oxidizing agents such as hydrogen peroxide or peracids. Examples 1 and 2 describe these mutants

which include F50/I124/Q222, F50/I124, F50/Q222, F50/L124/Q222, I124/Q222 and L124/Q222.

The third category of multiple subtilisin mutants comprises mutants with substitutions at position 222 combined with various substitutions at positions 166 or 169. These mutants, for example, combine the property of oxidative stability of the A222 mutation with the altered substrate specificity of the various 166 or 169 substitutions. Such multiple mutants include A166/A222, A166/C222, F166/C222, K166/A222, K166/C222, V166/A222 and V166/C222. The K166/A222 mutant subtilisin, for example, has a kcat/Km ratio which is approximately two times greater than that of the single A222 mutant subtilisin when compared using a substrate with phenylalanine as the P-1 amino acid. This category of multiple mutant is described in more detail in Example 12.

The fourth category of multiple mutants combines substitutions at position 156 (Glu to Q or S) with the substitution of Lys at position 166. Either of these single mutations improve enzyme performance upon substrates with glutamate as the P-1 amino acid. When these single mutations are combined, the resulting multiple enzyme mutants perform better than either precursor. See Example 9.

The fifth category of multiple mutants contain the substitution of up to four amino acids of the B. amyloliquefaciens subtilisin sequence. These mutants have specific properties which are virtually identicle to the properties of the subtilisin from B. licheniformis. The subtilisin from B. licheniformis differs from B. amyloliquefaciens subtilisin at 87 out of 275 amino acids. The multiple mutant

F50/S156/A169/L217 was found to have similar substrate specificity and kinetics to the licheniformis enzyme. (See Example 13.) However, this is probably due to only three of the mutations (S156, A169 and L217) which are present in the substrate binding region of the enzyme. It is quite surprising that, by making only three changes out of the 87 different amino acids between the sequence of the two enzymes, the B. amyloliquifaciens enzyme was converted into an enzyme with properties similar to B. licheniformis enzyme. Other enzymes in this series include F50/Q156/N166/L217 and F50/S156/L217.

The sixth category of multiple mutants includes the combination of substitutions at position 107 (Ile to V) with the substitution of Lys at position 213 with Arg, and the combination of substitutions of position 204 (preferably Ser to C or L but also to all other amino acids) with the substituion of Lys at position 213 with R. Other multiple mutants which have altered alkaline stability include Q156/K166, Q156/N166, S156/K166, S156/N166 (previously identified as having altered substrate specificity), and F50/S156/A169/L217 (previously identified as a mutant of B. amyloliquifaciens subtilisin having properties similar to subtilisin from B. licheniformis). The mutant F50/V107/R213 was constructed based on the observed increase in alkaline stability for the single mutants F50, V107 and R213. It was determined that the V107/R213 mutant had an increased alkaline stability as compared to the wild type subtilisin. In this particular mutant, the increased alkaline stability was the result of the cumulative stability of each of the individual mutations. Similarly, the mutant F50/V107/R213 had an even greater alkaline stability as compared to the V107/R213 mutant indicating that

the increase in the alkaline stability due to the F50 mutation was also cumulative.

Table IV summarizes the multiple mutants which have been made including those not mentioned above.

In addition, based in part on the above results, substitution at the following residues in subtilisin is expected to produce a multiple mutant having increased thermal and alkaline stability: Ser24, Met50, Ile107, Glu156, Gly166, Gly169, Ser204, Lys213, Gly215, and Tyr217.

TABLE IV

Double Mutants	Triple, Quadruple or Other Multiple
C22/C87	F50/I124/Q222
C24/C87	F50/L124/Q222
V45/V48	F50/L124/A222
C49/C94	A21/C22/C87
C49/C95	F50/S156/N166/L217
C50/C95	F50/Q156/N166/L217
C50/C110	F50/S156/A169/L217
F50/I124	F50/S156/L217
F50/Q222	F50/Q156/K166/L217
I124/Q222	F50/S156/K166/L217
Q156/D166	F50/Q156/K166/K217
Q156/K166	F50/S156/K166/K217
Q156/N166	F50/V107/R213
S156/D166	[S153/S156/A158/G159/S160/Δ161-164/I165/S166/A169/R170]
S156/K166	L204/R213
S156/N166	R213/204A, E, Q, D, N, G, K, V, R, T, P, I, M, F, Y, W or H
S156/A169	V107/R213
A166/A222	
A166/C223	
F166/A222	
F166/C222	
K166/A222	
K166/C222	
V166/A222	
V166/C222	
A169/A222	
A169/A222	
A169/C222	
A21/C22	

In addition to the above identified amino acid residues, other amino acid residues of subtilisin are

also considered to be important with regard to substrate specificity. Mutation of each of these residues is expected to produce changes in the substrate specificity of subtilisin. Moreover, multiple mutations among these residues and among the previously identified residues are also expected to produce subtilisin mutants having novel substrate specificity.

Particularly important residues are His67, Ile107, Leu126 and Leu135. Mutation of His67 should alter the S-1' subsite, thereby altering the specificity of the mutant for the P-1' substrate residue. Changes at this position could also affect the pH activity profile of the mutant. This residue was identified based on the inventor's substrate modeling from product inhibitor complexes.

Ile107 is involved in P-4 binding. Mutation at this position thus should alter specificity for the P-4 substrate residue in addition to the observed effect on alkaline stability. Ile107 was also identified by molecular modeling from product inhibitor complexes.

The S-2 binding site includes the Leu126 residue. Modification at this position should therefore affect P-2 specificity. Moreover, this residue is believed to be important to convert subtilisin to an amino peptidase. The pH activity profile should also be modified by appropriate substitution. These residues were identified from inspection of the refined model, the three dimensional structure from modeling studies. A longer side chain is expected to preclude binding of any side chain at the S-2 subsite. Therefore, binding would be restricted to subsites S-1, S-1', S-2', S-3'

and cleavage would be forced to occur after the amino terminal peptide.

Leu135 is in the S-4 subsite and if mutated should alter substrate specificity for P-4 if mutated. This residue was identified by inspection of the three-dimensional structure and modeling based on the product inhibitor complex of F222.

In addition to these sites, specific amino acid residues within the segments 97-103, 126-129 and 213-215 are also believed to be important to substrate binding.

Segments 97-103 and 126-129 form an antiparallel beta sheet with the main chain of substrate residues P-4 through P-2. Mutating residues in those regions should affect the substrate orientation through main chain (enzyme) - main chain (substrate) interactions, since the main chain of these substrate residues do not interact with these particular residues within the S-4 through S-2 subsites.

Within the segment 97-103, Gly97 and Asp99 may be mutated to alter the position of residues 101-103 within the segment. Changes at these sites must be compatible, however. In B. amyloliquifaciens subtilisin Asp99 stabilizes a turn in the main chain tertiary folding that affects the direction of residues 101-103. B. licheniformis subtilisin Asp97, functions in an analogous manner.

In addition to Gly97 and Asp99, Ser101 interacts with Asp99 in B. amyliquefaciens subtilisin to stabilize the same main chain turn. Alterations at this residue should alter the 101-103 main chain direction.

Mutations at Glu103 are also expected to affect the 101-103 main chain direction.

The side chain of Gly102 interacts with the substrate P-3 amino acid. Side chains of substituted amino acids thus are expected to significantly affect specificity for the P-3 substrate amino acids.

All the amino acids within the 127-129 segment are considered important to substrate specificity. Gly 127 is positioned such that its side chain interacts with the S-1 and S-3 subsites. Altering this residue thus should alter the specificity for P-1 and P-3 residues of the substrate.

The side chain of Gly128 comprises a part of both the S-2 and S-4 subsites. Altered specificity for P-2 and P-4 therefore would be expected upon mutation. Moreover, such mutation may convert subtilisin into an amino peptidase for the same reasons substitutions of Leu126 would be expected to produce that result.

The Pro129 residue is likely to restrict the conformational freedom of the sequence 126-133, residues which may play a major role in determining P-1 specificity. Replacing Pro may introduce more flexibility thereby broadening the range of binding capabilities of such mutants.

The side chain of Lys213 is located within the S-3 subsite. All of the amino acids within the 213-215 segment are also considered to be important to substrate specificity. Accordingly, altered P-3 substrate specificity is expected upon mutation of this residue.

The Tyr214 residue does not interact with substrate but is positioned such that it could affect the conformation of the hair pin loop 204-217.

Finally, mutation of the Gly215 residue should affect the S-3' subsite, and thereby alter P-3' specificity.

In addition to the above substitutions of amino acids, the insertion or deletion of one or more amino acids within the external loop comprising residues 152-172 may also affect specificity. This is because these residues may play a role in the "secondary contact region" described in the model of streptomyces subtilisin inhibitor complexed with subtilisin. Hirono, et al. (1984) J. Mol. Biol. 178, 389-413. Thermitase X has a deletion in this region, which eliminates several of these "secondary contact" residues. In particular, deletion of residues 161 through 164 is expected to produce a mutant subtilisin having modified substrate specificity. In addition, a rearrangement in this area induced by the deletion should alter the position of many residues involved in substrate binding, predominantly at P-1. This, in turn, should affect overall activity against proteinaceous substrates.

The effect of deletion of residues 161 through 164 has been shown by comparing the activity of the wild type (WT) enzyme with a mutant enzyme containing this deletion as well as multiple substitutions (i.e., S153/S156/A158/G159/S160/Δ161-164/I165/S166/A169/R170). This produced the following results:

TABLE V

	kcat	Km	kcat/Km
WT	50	1.4×10^{-4}	3.6×10^{5}
Deletion mutant	8	5.0×10^{-6}	1.6×10^{6}

The WT has a kcat 6 times greater than the deletion mutant but substrate binding is 28 fold tighter by the deletion mutant. The overall efficiency of the deletion mutant is thus 4.4 times higher than the WT enzyme.

All of these above identified residues which have yet to be substituted, deleted or inserted into are presented in Table VI.

TABLE VI

Substitution/Insertion/Deletion

Residues

His67	Ala152
Leu126	Ala153
Leu135	Gly154
Gly97	Asn155
Asp99	Gly156
Ser101	Gly157
Gly102	Gly160
Glu103	Thr158
Leu126	Ser159
Gly127	Ser161
Gly128	Ser162
Pro129	Ser163
Tyr214	Thr164
Gly215	Val165
Gly166	Gly169
Tyr167	Lys170
Pro168	Tyr171
	Pro172

The following disclosure is intended to serve as a representation of embodiments herein, and should not be construed as limiting the scope of this application. These specific examples disclose the construction of certain of the above identified mutants. The construction of the other mutants, however, is apparent from the disclosure herein and that presented in EPO Publication No. 0130756.

All literature citations are expressly incorporated by reference.

EXAMPLE 1

Identification of Peracid Oxidizable Residues of Subtilisin Q222 and L222

As shown in Figures 6A and 6B, organic peracid oxidants inactivate the mutant subtilisins Met222L and Met222Q (L222 and Q222). This example describes the identification of peracid oxidizable sites in these mutant subtilisins.

First, the type of amino acid involved in peracid oxidation was determined. Except under drastic conditions (Means, G.E., et al. (1971) Chemical Modifications of Proteins, Holden-Day, S.F., CA, pp. 160-162), organic peracids modify only methionine and tryptophan in subtilisin. Difference spectra of the enzyme over the 250nm to 350nm range were determined during an inactivation titration employing the reagent, diperdodecanoic acid (DPDA) as oxidant. Despite quantitative inactivation of the enzyme, no change in absorbance over this wavelength range was noted as shown in Figures 7A and 7B indicating that tryptophan was not oxidized. Fontana, A., et al. (1980) Methods in Peptide and Protein Sequence

Analysis (C. Birr ed.) Elsevier, New York, p. 309. The absence of tryptophan modification implied oxidation of one or more of the remaining methionines of B. amyloliquefaciens subtilisin. See Figure 1.

To confirm this result the recombinant subtilisin Met222F was cleaved with cyanogen bromide (CNBr) both before and after oxidation by DPDA. The peptides produced by CNBr cleavage were analyzed on high resolution SDS-pyridine peptide gels (SPG).

Subtilisin Met222F (F222) was oxidized in the following manner. Purified F222 was resuspended in 0.1 M sodium borate pH 9.5 at 10 mg/ml and was added to a final concentration of 26 diperdodecanoic acid (DPDA) at 26 mg/ml was added to produce an effective active oxygen concentration of 30 ppm. The sample was incubated for at least 30 minutes at room temperature and then quenched with 0.1 volume of 1 M Tris pH 8.6 buffer to produce a final concentration of 0.1 M Tris pH 8.6). 3mM phenylmethylsulfonyl fluoride (PMSF) was added and 2.5 ml of the sample was applied to a Pharmacia PD10 column equilibrated in 10 mM sodium phosphate pH 6.2, 1 mM PMSF. 3.5 ml of 10 mM sodium phosphate pH6.2, 1mM PMSF was applied and the eluant collected.

F222 and DPDA oxidized F222 were precipitated with 9 volumes of acetone at -20°C. The samples were resuspended at 10 mg/ml in 8M urea in 88% formic acid and allowed to sit for 5 minutes. An equal volume of 200 mg/ml CNBr in 88% formic acid was added (5 mg/ml protein) and the samples incubated for 2 hours at room temperature in the dark. Prior to gel electrophoresis, the samples were lyophilized and resuspended at 2-5 mg/ml in sample buffer (1%

pyridine, 5% $NaDodSO_4$, 5% glycerol and bromophenol blue) and disassociated at 95°C for 3 minutes.

The samples were electrophoresed on discontinuous polyacrylamide gels (Kyte, J., et al. (1983) Anal. Bioch. 133, 515-522). The gels were stained using the Pharmacia silver staining technique (Sammons, D.W., et al. (1981) Electrophoresis 2 135-141).

The results of this experiment are shown in Figure 8. As can be seen, F222 treated with CNBr only gives nine resolved bands on SPG. However, when F222 is also treated with DPDA prior to cleavage, bands X, 7 and 9 disappear whereas bands 5 and 6 are greatly increased in intensity.

In order to determine which of the methionines were effected, each of the CNBr peptides was isolated by reversed phase HPLC and further characterized. The buffer system in both Solvent A (aqueous) and Solvent B (organic) for all HPLC separations was 0.05% triethylamime/trifloroacetic acid (TEA-TFA). In all cases unless noted, solvent A consisted of 0.05% TEA-TFA in H_2O, solvent B was 0.05% TEA-TFA in 1-propanol, and the flow rate was 0.5 ml/minute.

For HPLC analysis, two injections of 1 mg enzyme digest were used. Three samples were acetone precipitated, washed and dried. The dried 1 mg samples were resuspended at 10 mg/ml in 8M urea, 88% formic acid; an equal volume of 200 mg/ml CNBr in 88% formic acid was added (5 mg/ml protein). After incubation for 2 hours in the dark at room temperature, the samples were desalted on a 0.8 cm X 7

cm column of Tris Acryl GF05 coarse resin (IBF, Paris, France) equilibrated with 40% solvent B, 60% solvent A. 200 ul samples were applied at a flow rate of 1 ml a minute and 1.0-1.2 ml collected by monitoring the absorbance at 280nm. Prior to injection on the HPLC, each desalted sample was diluted with 3 volumes of solvent A. The samples were injected at 1.0 ml/min (2 minutes) and the flow then adjusted to 0.5 ml/min (100% A). After 2 minutes, a linear gradient to 60% B at 1.0% B/min was initiated. From each 1 mg run, the pooled peaks were sampled (50ul) and analyzed by gel electrophoresis as described above.

Each polypeptide isolated by reversed phase HPLC was further analyzed for homogeneity by SPG. The position of each peptide on the known gene sequence (Wells, J.A., et al. (1983) Nucleic Acids Res. 11 7911-7924) was obtained through a combination of amino acid compositional analysis and, where needed, amino terminal sequencing.

Prior to such analysis the following peptides were to rechromatographed.

1. CNBr peptides from F222 not treated with DPDA:

Peptide 5 was subjected to two additional reversed phase separations. The 10 cm C4 column was equilibrated to 80%A/ 20%B and the pooled sample applied and washed for 2 minutes. Next an 0.5% ml B/min gradient was initiated. Fractions from this separation were again rerun, this time on the 25 cm C4 column, and employing 0.05% TEA-TPA in acetonitrile/1-propanol (1:1) for solvent B. The gradient was identical to the one just described.

Peptide "X" was subjected to one additional separation after the initial chromatography. The sample was applied and washed for 2 minutes at 0.5ml/min (100%A), and a 0.5% ml B/min gradient was initiated.

Peptides 7 and 9 were rechromatographed in a similar manner to the first rerun of peptide 5.

Peptide 8 was purified to homogeneity after the initial separation.

2. CNBr Peptides from DPDA Oxidized F222:

Peptides 5 and 6 from a CNBr digest of the oxidized F222 were purified in the same manner as peptide 5 from the untreated enzyme.

Amino acid compositional analysis was obtained as follows. Samples (~1nM each amino acid) were dried, hydrolyzed in vacuo with 100 ul 6N HCl at 106°C for 24 hours and then dried in a Speed Vac. The samples were analyzed on a Beckmann 6300 AA analyzer employing ninhydrin detection.

Amino terminal sequence data was obtained as previously described (Rodriguez, N., et al. (1984) Anal. Biochem. 134, 538-547).

The results are shown in Table VII and Figure 9.

TABLE VII

Amino and COOH terminii of CNBr fragments

	Terminus and Method	
Fragment	amino, method	COOH, method
X	1, sequence	50, composition
9	51, sequence	119, composition
7	125, sequence	199, composition
8	200, sequence	275, composition
5ox	1, sequence	119, composition
6ox	120, composition	199, composition

Peptides 5ox and 6ox refer to peptides 5 and 6 isolated from CNBr digests of the oxidized protein where their respective levels are enhanced.

From the data in Table VII and the comparison of SPG tracks for the oxidized and native protein digests in Figure 8, it is apparent that (1) Met50 is oxidized leading to the loss of peptides X and 9 and the appearance of 5; and (2) Met124 is also oxidized leading to the loss of peptide 7 and the accumulation of peptide 6. Thus oxidation of B. amyloliquifaciens subtilisin with the peracid, diperdodecanoic acid leads to the specific oxidation of methionine at residues 50 and 124.

EXAMPLE 2

Substitution at Met50 and Met124 in Subtilisin Met222Q

The choice of amino acid for substitution at Met50 was based on the available sequence data for subtilisins

from B. licheniformis (Smith, E.C., et al. (1968) J. Biol. Chem. 243, 2184-2191), B.DY (Nedkov, P., et al. (1983) Hoppe Sayler's Z. Physiol. Chem. 364 1537-1540), B. amylosacchariticus (Markland, F.S., et al. (1967) J. Biol. Chem. 242 5198-5211) and B. subtilis (Stahl, M.L., et al. (1984) J. Bacteriol. 158, 411-418). In all cases, position 50 is a phenylalanine. See Figure 5. Therefore, Phe50 was chosen for construction.

At position 124, all known subtilisins possess a methionine. See Figure 5. Molecular modelling of the x-ray derived protein structure was therefore required to determine the most probable candidates for substitution. From all 19 candidates, isoleucine and leucine were chosen as the best residues to employ. In order to test whether or not modification at one site but not both was sufficient to increase oxidative stability, all possible combinations were built on the Q222 backbone (F50/Q222, I124/Q222, F50/I124/Q222).

A. Construction of Mutations Between Codons 45 and 50

All manipulations for cassette mutagenesis were carried out on pS4.5 using methods disclosed in EPO Publication No. 0130756 and Wells, J.A., et al, (1985) Gene 34, 315-323. The pA50 in Fig. 10, line 4, mutations was produced using the mutagenesis primer shown in Fig. 10, line 6, and employed an approach designated as restriction-purification which is described below. Briefly, a M13 template containing the subtilisin gene, M13mp11-SUBT was used for heteroduplex synthesis (Adelman, et al (1983), DNA 2, 183-193). Following transfection of JM101 (ATCC 33876), the 1.5 kb EcoRI-BamHI fragment containing the

subtilisin gene was subcloned from M13mp11 SUBT rf into a recipient vector fragment of pBS42 the construction of which is described in EPO Publication No. 0130756. To enrich for the mutant sequence (pΔ50, line 4), the resulting plasmid pool was digested with KpnI, and linear molecules were purified by polyacrylamide gel electrophoresis. Linear molecules were ligated back to a circular form, and transformed into E. coli KM294 cells (ATCC 31446). Isolated plasmids were screened by restriction analysis for the KpnI site. KpnI$^+$ plasmids were sequenced and confirmed the pΔ50 sequence. Asterisks in Figure 11 indicate the bases that are mutated from the wid type sequence (line 4). pΔ50 (line 4) was cut with StuI and EcoRI and the 0.5 Kb fragment containing the 5' half of the subtilisin gene was purified (fragment 1). pΔ50 (line 4) was digested with KpnI and EcoRI and the 4.0 Kb fragment containing the 3' half of the subtilisin gene and vector sequences was purified (fragment 2). Fragments 1 and 2 (line 5), and duplex DNA cassettes coding for mutations desired (shaded sequence, line 6) were mixed in a molar ratio of 1:1:10, respectively. For the particular construction of this example the DNA cassette contained the triplet TTT for codon 50 which encodes Phe. This plasmid was designated pF50. The mutant subtilisin was designated F50.

B. Construction of Mutation Between Codons 122 and 127

The procedure of Example 2A was followed in substantial detail except that the mutagenesis primer of Figure 11, line 7 was used and restriction-purification for the EcoRV site in pΔ124 was used. In addition, the DNA cassette (shaded sequence, Figure

11, line 6) contained the triplet ATT for codon 124 which encodes Ile and CTT for Leu. Those plasmids which contained the substitution of Ile for Met124were designeated pI124. The mutant subtilisin was designated I124.

C. Construction of Various F50/I124/Q222 Multiple Mutants

The triple mutant, F50/I124/Q222, was constructed from a three-way ligation in which each fragment contained one of the three mutations. The single mutant Q222 (pQ222) was prepared by cassette mutagenesis as described in EPO Publication No. 0130756. The F50 mutation was contained on a 2.2Kb AvaII to PvuII fragment from pF50; the I124 mutation was contained on a 260 bp PvuII to AvaII fragment from pI124; and the Q222 mutation was contained on 2.7 kb AvaII to AvaII fragment from pQ222. The three fragments were ligated together and transformed into E. coli MM294 cells. Restriction analysis of plasmids from isolated transformants confirmed the construction. To analyze the final construction it was convenient that the AvaII site at position 798 in the wild-type subtilisin gene was eliminated by the I124 construction.

The F50/Q222 and I124/Q222 mutants were constructed in a similar manner except that the appropriate fragment from pS4.5 was used for the final construction.

D. Oxidative Stability of Q222 Mutants

The above mutants were analyzed for stability to peracid oxidation. As shown in Fig. 12, upon incubation with diperdodecanoic acid (protein 2mg/mL, oxidant 75ppm[O]), both the I124/Q222 and the

F50/I124/Q222 are completely stable whereas the F50/Q222 and the Q222 are inactivated. This indicates that conversion of Met124 to I124 in subtilisin Q222 is sufficient to confer resistance to organic peracid oxidants.

EXAMPLE 3

Subtilisin Mutants Having Altered Substrate Specificity-Hydrophobic Substitutions at Residues 166

Subtilisin contains an extended binding cleft which is hydrophobic in character. A conserved glycine at residue 166 was replaced with twelve non-ionic amino acids which can project their side-chains into the S-1 subsite. These mutants were constructed to determine the effect of changes in size and hydrophobicity on the binding of various substrates.

A. Kinetics for Hydrolysis of Substrates Having Altered P-1 Amino Acids by Subtilisin from B. Amyloliquefaciens

Wild-type subtilisin was purified from B. subtilis culture supernatants expressing the B. amyloliquefaciens subtilisin gene (Wells, J.A., et al. (1983) Nucleic Acids Res. 11, 7911-7925) as previously described (Estell, D.A., et al. (1985) J. Biol. Chem. 260, 6518-6521). Details of the synthesis of tetrapeptide substrates having the form succinyl-L-AlaL-AlaL-ProL-[X]-p-nitroanilide (where X is the P1 amino acid) are described by DelMar, E.G., et al. (1979) Anal. Biochem. 99, 316-320. Kinetic parameters, Km(M) and kcat(s^{-1}) were measured using a modified progress curve analysis (Estell, D.A., et al. (1985) J. Biol. Chem. 260, 6518-6521). Briefly, plots

of rate versus product concentration were fit to the differential form of the rate equation using a non-linear regression algorithm. Errors in kcat and Km for all values reported are less than five percent. The various substrates in Table VIII are ranged in order of decreasing hydrophobicity. Nozaki, Y. (1971), J. Biol. Chem. 246, 2211-2217; Tanford C. (1978) Science 200, 1012).

TABLE VIII

P1 substrate Amino Acid	kcat(S^{-1})	1/Km(M^{-1})	kcat/Km ($s^{-1}M^{-1}$)
Phe	50	7,100	360,000
Tyr	28	40,000	1,100,000
Leu	24	3,100	75,000
Met	13	9,400	120,000
His	7.9	1,600	13,000
Ala	1.9	5,500	11,000
Gly	0.003	8,300	21
Gln	3.2	2,200	7,100
Ser	2.8	1,500	4,200
Glu	0.54	32	16

The ratio of kcat/Km (also referred to as catalytic efficienty) is the apparent second order rate constant for the conversion of free enzyme plus substrate (E+S) to enzyme plus products (E+P) (Jencks, W.P., Catalysis in Chemistry and Enzymology (McGraw-Hill, 1969) pp. 321-436; Fersht, A., Enzyme Structure and Mechanism (Freeman, San Francisco, 1977) pp. 226-287). The log (kcat/Km) is proportional to transition state binding

energy, $\Delta G_T^{\ddagger}$. A plot of the log kcat/Km versus the hydrophobicity of the P1 side-chain (Figure 14) shows a strong correlation (r = 0.98), with the exception of the glycine substrate which shows evidence for non-productive binding. These data show that relative differences between transition-state binding energies can be accounted for by differences in P-1 side-chain hydrophobicity. When the transition-state binding energies are calculated for these substrates and plotted versus their respective side-chain hydrophobicities, the line slope is 1.2 (not shown). A slope greater than unity, as is also the case for chymotrypsin (Fersht, A., Enzyme Structure and Mechanism (Freeman, San Francisco, 1977) pp. 226-287; Harper, J.W., et al. (1984) Biochemistry, 23, 2995-3002), suggests that the P1 binding cleft is more hydrophobic than ethanol or dioxane solvents that were used to empirically determine the hydrophobicity of amino acids (Nozaki, Y., et al. J. Biol. Chem. (1971) 246, 2211-2217; Tanford, C. (1978) Science 200, 1012).

For amide hydrolysis by subtilisin, kcat can be interpreted as the acylation rate constant and Km as the dissociation constant, for the Michaelis complex (E·S), Ks. Gutfreund, H., et al (1956) Biochem. J. 63, 656. The fact that the log kcat, as well as log 1/Km, correlates with substrate hydrophobicity is consistent with proposals (Robertus, J.D., et al. (1972) Biochemistry 11, 2439-2449; Robertus, J.D., et al. (1972) Biochemistry 11, 4293-4303) that during the acylation step the P-1 side-chain moves deeper into the hydrophobic cleft as the substrate advances from the Michaelis complex (E·S) to the tetrahedral transition-state complex ($E \cdot S^{\ddagger}$). However, these data can also be interpreted as the hydrophobicity of the P1 side-chain effecting the orientation, and thus the

susceptibility of the scissile peptide bond to nucleophilic attack by the hydroxyl group of the catalytic Ser221.

The dependence of kcat/Km on P-1 side chain hydrophobicity suggested that the kcat/Km for hydrophobic substrates may be increased by increasing the hydrophobicity of the S-1 binding subsite. To test this hypothesis, hydrophobic amino acid substitutions of Gly166 were produced.

Since hydrophobicity of aliphatic side-chains is directly proportional to side-chain surface area (Rose, G.D., et al. (1985) Science 229, 834-838; Reynolds, J.A., et al. (1974) Proc. Natl. Acad. Sci. USA 71, 2825-2927), increasing the hydrophobicity in the S-1 subsite may also sterically hinder binding of larger substrates. Because of difficulties in predicting the relative importance of these two opposing effects, we elected to generate twelve non-charged mutations at position 166 to determine the resulting specificities against non-charged substrates of varied size and hydrophobicity.

B. Cassette Mutagenesis of the P1 Binding Cleft

The preparation of mutant subtilisims containing the substitution of the hydrophobic amino acids Ala, Val and Phe into residue 166 has been described in EPO Publication No. 0130756. The same method was used to produce the remaining hydrophobic mutants at residue 166. In applying this method, two unique and silent restriction sites were introduced in the subtilisin genes to closely flank the target codon 166. As can be seen in Figure 13, the wild type sequence (line 1)

was altered by site-directed mutagenesis in M13 using the indicated 37mer mutagenesis primer, to introduce a 13 bp delection (dashedline) and unique SacI and XmaI sites (underlined sequences) that closely flank codon 166. The subtilisin gene fragment was subcloned back into the E. coli - B. subtilis shuttle plasmid, pBS42, giving the plasmid pΔ166 (Figure 13, line 2). pΔ166 was cut open with SacI and XmaI, and gapped linear molecules were purified (Figure 13, line 3). Pools of synthetic oligonucleotides containing the mutation of interest were annealed to give duplex DNA cassettes that were ligated into gapped pΔ166 (underlined and overlined sequences in Figure 13, line 4). This construction restored the coding sequence except over position 166(NNN; line 4). Mutant sequences were confirmed by dideoxy sequencing. Asterisks denote sequence changes from the wild type sequence. Plasmids containing each mutant B. amyloliquefaciens subtilisin gene were expressed at roughly equivalent levels in a protease deficient strain of B. subtilis, BG2036 as previously described. EPO Publication No. 0130756; Yang, M., et al. (1984) J. Bacteriol. 160, 15-21; Estell, D.A., et al (1985) J. Biol. Chem. 260, 6518-6521.

C. Narrowing Substrate Specificity by Steric Hindrance

To probe the change in substrate specificity caused by steric alterations in the S-1 subsite, position 166 mutants were kinetically analyzed versus P1 substrates of increasing size (i.e., Ala, Met, Phe and Tyr). Ratios of kcat/Km are presented in log form in Figure 15 to allow direct comparisons of transition-state binding energies between various enzyme-substrate pairs.

According to transition state theory, the free enery difference between the free enzyme plus substrate (E + S) and the transition state complex ($E \cdot S^{\ddagger}$) can be calculated from equation (1),

(1) $$\Delta G_T^{\ddagger} = -RT \ln kcat/Km + RT \ln kT/h$$

in which kcat is the turnover number, Km is the Michaelis constant, R is the gas constant, T is the temperature, k is Boltzmann's constant, and h is Planck's constant. Specificity differences are ezpressed quantitatively as differences between transition state binding energies (i.e., $\Delta\Delta G_t^{\ddagger}$), and can be calculated from equation (2).

(2) $$\Delta\Delta G_T^{\ddagger} = -RT \ln (kcat/Km)_A/(kcat/Km)_B$$

A and B represent either two different substrates assayed againt the same enzyme, or two mutant enzymes assayed against the same substrate.

As can be seen from Figure 15A, as the size of the side-chain at position 166 increases the substrate preference shifts from large to small P-1 side-chains. Enlarging the side-chain at position 166 causes kcat/Km to decrease in proportion to the size of the P-1 substrate side-chain (e.g., from Gly166 (wild-type) through W166, the kcat/Km for the Tyr substrate is decreased most followed in order by the Phe, Met and Ala P-1 substrates).

Specific steric changes in the position 166 side-chain, such as he presence of a β-hydroxyl group, β- or γ-aliphatic branching, cause large decreases in kcat/Km for larger P1 substrates. Introducing a β-hydroxyl group in going from A166 (Figure 15A) to

S166 (Figure 15B), causes an 8 fold and 4 fold reduction in kcat/Km for Phe and Tyr substrates, respectively, while the values for Ala and Met substrates are unchanged. Producing a β-branched structure, in going from S166 to T166, results in a drop of 14 and 4 fold in kcat/Km for Phe and Tyr, respectively. These differences are slightly magnified for V166 which is slightly larger and isosteric with T166. Enlarging the β-branched substituents from V166 to I166 causes a lowering of kcat/Km between two and six fold toward Met, Phe and Tyr substrates. Inserting a γ-branched structure, by replacing M166 (Figure 15A) with L166 (Figure 15B), produces a 5 fold and 18 fold decrease in kcat/Km for Phe and Tyr substrates, respectively. Aliphatic γ-branched appears to induce less steric hindrance toward the Phe P-1 substrate than β-branching, as evidenced by the 100 fold decrease in kcat/Km for the Phe substrate in going from L166 to I166.

Reductions in kcat/Km resulting from increases in side chain size in the S-1 subsite, or specific structural features such as β- and γ-branching, are quantitatively illustrated in Figure 16. The kcat/Km values for the position 166 mutants determined for the Ala, Met, Phe, and Tyr P-1 substrates (top panel through bottom panel, respectively), are plotted versus the position 166 side-chain volumes (Chothia, C. (1984) Ann. Rev. Biochem. 53, 537-572). Catalytic efficiency for the Ala substrate reaches a maximum for I166, and for the Met substrate it reaches a maximum between V166 and L166. The Phe substrate shows a broad kcat/Km peak but is optimal with A166. Here, the β-branched position 166 substitutions form a line that is parallel to, but roughly 50 fold lower in kcat/Km than side-chains of similar size [i.e., C166 versus

T166, L166 versus I166]. The Tyr substrate is most efficiently utilized by wild type enzyme (Gly166), and there is a steady decrease as one proceeds to large position 166 side-chains. The β-branched and γ-branched substitutions form a parallel line below the other non-charged substitutions of similar molecular volume.

The optimal substitution at position 166 decreases in volume with increasing volume of the P1 substrate [i.e., I166/Ala substrate, L166/Met substrate, A166/Phe substrate, Gly166/Tyr substrate]. The combined volumes for these optimal pairs may approximate the volume for productive binding in the S-1 subsite. For the optimal pairs, Gly166/Tyr substrate, A166/Phe substrate, L166/Met substrate, V166/Met substrate, and I166/Ala substrate, the combined volumes are 266,295,313,339 and 261 A^3, respectively. Subtracting the volume of the peptide backbone from each pair (i.e., two times the volume of glycine), an average side-chain volume of 160±32A^3 for productive binding can be calculated.

The effect of volume, in excess to the productive binding volume, on the drop in transition-state binding energy can be estimated from the Tyr substrate curve (bottom panel, Figure 16), because these data, and modeling studies (Figure 2), suggest that any substitution beyond glycine causes steric repulsion. A best-fit line drawn to all the data ($r = 0.87$) gives a slope indicating a loss of roughly 3 kcal/mol in transition state binding energy per 100A^3 of excess volume. (100A^3 is approximately the size of a leucyl side-chain.)

D. Enhanced Catalytic Efficiency Correlates with Increasing Hydrophobicity of the Position 166 Substitution

Substantial increases in kcat/Km occur with enlargement of the position 166 side-chain, except for the Tyr P-1 substrate (Figure 16). For example, kcat/Km increases in progressing from Gly166 to I166 for the Ala substrate (net of ten-fold), from Gly166 to L166 for the Met substrate (net of ten-fold) and from Gly166 to A166 for the Phe substrate (net of two-fold). The increases in kcat/Km cannot be entirely explained by the attractive terms in the van der Waals potential energy function because of their strong distance dependance ($1/r^6$) and because of the weak nature of these attractive forces (Jencks, W.P., Catalysis in Chemistry and Enzymology (McGraw-Hill, 1969) pp. 321-436; Fersht, A., Enzyme Structure and Mechanism (Freeman, San Francisco, 1977) pp. 226-287; Levitt, M. (1976) J. Mol. Biol. 104, 59-107). For example, Levitt (Levitt, M. (1976) J. Mol. Biol. 104, 59-107) has calculated that the van der Waals attraction between two methionyl residues would produce a maximal interaction energy of roughly -0.2 kcal/mol. This energy would translate to only 1.4 fold increase in kcat/Km.

The increases of catalytic efficiency caused by side-chain substitutions at position 166 are better accounted for by increases in the hydrophobicity of the S-1 subsite. The increase kcat/Km observed for the Ala and Met substrates with increasing position 166 side-chain size would be expected, because hydrophobicity is roughly proportional to side-chain surface area (Rose, G.D., et al. (1985) Science 229, 834-838; Reynolds, J.A., et al. (1974) Proc. Natl. Acad. Sci. USA 71, 2825-2927).

Another example that can be interpreted as a hydrophobic effect is seen when comparing kcat/Km for isosteric substitutions that differ in hydrophobicity such as S166 and C166 (Figure 16). Cysteine is considerably more hydrophobic than serine (-1.0 versus +0.3 kcal/mol) (Nozaki, Y., et al. (1971) J. Biol. Chem. 246, 2211-2217; Tanford, C. (1978) Science 200, 1012). The difference in hydrophobicity correlates with the observation that C166 becomes more efficient relative to Ser166 as the hydrophobicity of the substrates increases (i.e., Ala < Met < Tye < Phe). Steric hindrance cannot explain these differences because serine is considerably smaller than cysteine (99 versus 118A^3). Paul, I.C., Chemistry of the -SH Group (ed. S. Patai, Wiley Interscience, New York, 1974) pp. 111-149.

E. Production of an Elastase-Like Specificity in Subtilisin

The I166 mutation illustrates particularly well that large changes in specificity can be produced by altering the structure and hydrophobicity of the S-1 subsite by a single mutation (Figure 17). Progressing through the small hydrophobic substrates, a maximal specificity improvement over wild type occurs for the Val substrate (16 fold in kcat/Km). As the substrate side chain size increases, these enhancements shrink to near unity (i.e., Leu and His substrates). The I166 enzyme becomes poorer against larger aromatic substrates of increasing size (e.g., I166 is over 1,000 fold worse against the Tyr substrate than is Gly166). We interpret the increase in catalytic efficiency toward the small hydrophobic substrates for I166 compared to Gly166 to the greater hydrophobicity of isoluecine (i.e., -1.8 kcal/mol versus 0). Nozaki,

Y., et al. (1971) J. Biol. Chem. 246, 2211-2217; Tanford, C. (1978) Science 200, 1012. The decrease in catalytic efficiency toward the very large substrates for I166 versus Gly166 is attributed to steric repulsion.

The specificity differences between Gly166 and I166 are similar to the specificity differences between chymotrypsin and the evolutionary relative, elastase (Harper, J.W., et al (1984) Biochemistry 23, 2995-3002). In elastase, the bulky amino acids, Thr and Val, block access to the P-1 binding site for large hydrophobic substrates that are preferred by chymotrypsin. In addition, the catalytic efficiencies toward small hydrophobic substrates are greater for elastase than for chymotrypsin as we obeseve for I166 versus Gly166 in subtilisin.

EXAMPLE 4

Substitution of Ionic Amino Acids for Gly166

The construction of subtilisin mutants containing the substitution of the ionic amino acids Asp, Asn, Gln, Lys and Ang are disclosed in EPO Publication No. 0130756. The present example describes the construction of the mutant subtilisin containing Glu at position 166 (E166) and presents substrate specificity data on these mutants. Further data on position 166 and 156 single and double mutants is presented infra.

pΔ166, described in Example 3, was digested with SacI and XmaI. The double strand DNA cassette (underlined and overlined) of line 4 in Figure 13 contained the

triplet GAA for the codon 166 to encode the replacement of Glu for Gly166. This mutant plasmid designated pQ166 was propagated in BG2036 as described. This mutant subtilisin, together with the other mutants containing ionic substituent amino acids at residue 166, were isolated as described and further analyzed for variations in substrate specificity.

Each of these mutants was analyzed with the tetrapeptide substrates, succinyl-L-AlaL-AlaProL-X--p-nitroanilide, where X was Phe, Ala and Glu.

The results of this analysis are shown in Table IX.

TABLE IX

	P-1 Substrate (kcat/Km x 10^{-4})		
Position 166	Phe	Ala	Glu
Gly (wild type)	36.0	1.4	0.002
Asp (D)	0.5	0.4	<0.001
Glu (E)	3.5	0.4	<0.001
Asn (N)	18.0	1.2	0.004
Gln (Q)	57.0	2.6	0.002
Lys (K)	52.0	2.8	1.2
Arg (R)	42.0	5.0	0.08

These results indicate that charged amino acid substitutions at Gly166 have improved catalytic efficiencies (kcat/Km) for oppositely charged P-1 substrates (as much as 500 fold) and poorer catalytic efficiency for like charged P-1 substrates.

EXAMPLE 5

Substitution of Glycine at Position 169

The substitution of Gly169 in B. amyloliquefaciens subtilisin with Ala and Ser is described in EPO Publication No. 0130756. The same method was used to make the remaining 17 mutants containing all other substituent amino acids for position 169.

The construction protocol is summarized in Figure 18. The overscored and underscored double stranded DNA cassettes used contained the following triplet encoding the substitution of the indicated amino acid at residue 169.

GCT	A	ATG	M
TGT	C	AAC	N
GAT	D	CCT	P
GAA	E	CAA	Q
TTC	F	AGA	R
GGC	G	AGC	S
CAC	H	ACA	T
ATC	I	GTT	V
AAA	K	TGG	W
CTT	L	TAC	Y

Each of the plasmids containing a substituted Gly169 was designated pX169, where X represents the substituent amino acid. The mutant subtilisins were simialrly designated.

Two of the above mutant subtilisins, A169 and S169, were analyzed for substrate specificity against synthetic substrates containing Phe, Leu, Ala and Arg in the P-1 position. The following results are shown in Table X.

TABLE X

Effect of Serine and Alanine Mutations at Position 169 on P-1 Substrate Specificity

	P-1 Substrate (kcat/Km x 10^{-4})			
Position 169	Phe	Leu	Ala	Arg
Gly (wild type)	40	10	1	0.4
Al69	120	20	1	0.9
S169	50	10	1	0.6

These results indicate that substitutions of Ala and Ser at Gly169 have remarkably similar catalytic efficiencies against a range of P-1 substrates compared to their position 166 counterparts. This is probably because position 169 is at the bottom of the P-1 specificity subsite.

EXAMPLE 6

Substitution at Position 104

Tyr104 has been substituted with Ala, His, Leu, Met and Ser. The method used was a modification of the site directed mutagenesis method. According to the protocol of Figure 19, a primer (shaded in line 4) introduced a unique HindIII site and a frame shift mutation at codon 104. Restriction-purification for the unique HindIII site facilitated the isolation of the mutant sequence (line 4). Restriction-selection against this HindIII site using pimers in line 5 was used to obtain position 104 mutants.

The following triplets were used in the primers of Figure 19, line 5 for the 104 codon which substituted the following amino acids.

GCT	A	TTC	F
ATG	M	CCT	P
CTT	L	ACA	T
AGC	S	TGG	W
CAC	H	TAC	Y
CAA	Q	GTT	V
GAA	E	AGA	R
GGC	G	AAC	N
ATC	I	GAT	D
AAA	K	TGT	C

The substrates in Table XI were used to analyze the substrate specificity of these mutants. The results obtained fo H104 subtilisin are shown in Table XI.

TABLE XI

	kcat		Km		Kcat/Km	
Substrate	WT	H104	WT	H104	WT	H104
sAAPFpNA	50.0	22.0	1.4×10^{-4}	7.1×10^{-4}	3.6×10^{5}	3.1×10^{4}
sAAPApNA	3.2	2.0	2.3×10^{-4}	1.9×10^{-3}	1.4×10^{4}	1×10^{3}
sFAPFpNA	26.0	38.0	1.8×10^{-4}	4.1×10^{-4}	1.5×10^{5}	9.1×10^{4}
sFAPApNA	0.32	2.4	7.3×10^{-5}	1.5×10^{-4}	4.4×10^{3}	1.6×10^{4}

From these data it is clear that the substitution of His for Tyr at position 104 produces an enzyme which is more efficient (higher kcat/Km) when Phe is at the P-4 substrate position than when Ala is at the P-4 substrate position.

EXAMPLE 7

Substitution of Ala152

Ala152 has been substituted by Gly and Ser to determine the effect of such substitutions on substrate specificity.

The wild type DNA sequence was mutated by the V152/P153 primer (Figure 20, line 4) using the above restriction-purification approach for the new KpnI site. Other mutant primers (shaded sequences Figure 20; S152, line 5 and G152, line 6) mutated the new KpnI site away and such mutants were isolated using the restriction-selection procedure as described above for loss of the KpnI site.

The results of these substitutions for the above synthetic substrates containing the P-1 amino acids Phe, Leu and Ala are shown in Table XII.

TABLE XII

	P-1 Substrate ($kcat/Kmx10^{-4}$)		
Position 152	Phe	Leu	Ala
Gly (G)	0.2	0.4	<0.04
Ala (wild type)	40.0	10.0	1.0
Ser (S)	1.0	0.5	0.2

These results indicate that, in contrast to positions 166 and 169, replacement of Ala152 with Ser or Gly causes a dramatic reduction in catalytic efficiencies

across all substrates tested. This suggests Ala152, at the top of the S-1 subsite, may be the optimal amino acid because Ser and Gly are homologous Ala substitutes.

EXAMPLE 8

Substitution at Position 156

Mutants containing the substitution of Ser and Gln for Glu156 have been constructed according to the overall method depicted in Figure 21. This method was designed to facilitate the construciton of multiple mutants at position 156 and 166 as will be described hereinafter. However, by regenerating the wild type Gly166, single mutations at Glu156 were obtained.

The plasmid pΔ166 is already depicted in line 2 of Figure 13. The synthetic oligonucleotides at the top right of Figure 21 represent the same DNA cassettes depicted in line 4 of Figure 13. The plasmid p166 in Figure 21 thus represents the mutant plasmids of Examples 3 and 4. In this particular example, p166 contains the wild type Gly166.

Construction of position 156 single mutants were prepared by ligation of the three fragments (1-3) indicated at the bottom of Figure 21. Fragment 3, containing the carboxy-terminal portion of the subtilisin gene including the wild type position 166 codon, was isolated as a 610 bp SacI-BamHI fragment. Fragment 1 contained the vector sequences, as well as the amino-terminal sequences of the subtilisin gene through codon 151. To produce fragment 1, a unique KpnI site at codon 152 was introduced into the wild type subtilisin sequence from pS4.5. Site-directed

mutagenesis in M13 employed a primer having the sequence 5'-TA-GTC-GTT-GCG-GTA-CCC-GGT-AAC-GAA-3' to produce the mutation. Enrichment for the mutant sequence was accomplished by restriction with KpnI, purification and self ligation. The mutant sequence containing the KpnI site was confirmed by direct plasmid sequencing to give pV152. pV152 (~1 μg) was digested with KpnI and treated with 2 units of DNA polymerase I large fragment (Klenow fragment from Boeringer-Mannheim) plus 50 μM deoxynucleotide triphosphates at 37°C for 30 min. This created a blunt end that terminated with codon 151. The DNA was extracted with 1:1 volumes phenol and $CHCl_3$ and DNA in the aqueous phase was precipitated by addition of 0.1 volumes 5M ammonium acetate and two volumes ethanol. After centrifugation and washing the DNA pellet with 70% ethanol, the DNA was lyophilized. DNA was digested with BamHI and the 4.6kb piece (fragment 1) was purified by acrylamide gel electrophoresis followed by electroelution. Fragment 2 was a duplex synthetic DNA cassette which when ligated with fragments 1 and 3 properly restored the coding sequence except at codon 156. The top strand was synthesized to contain a glutamine codon, and the complementary bottom strand coded for serine at 156. Ligation of heterophosphorylated cassettes leads to a large and favorable bias for the phosphorylated over the non-phosphorylated oligonucleotide sequence in the final segrated plasmid product. Therefore, to obtain Q156 the top strand was phosphorylated, and annealed to the non-phosphorylated bottom strand prior to ligation. Similarly, to obtain S156 the bottom strand was phosphorylated and annealed to the non-phosphorylated top strand. Mutant sequences were isolated after ligation and transformation, and were confirmed by restriction analysis and DNA sequencing

as before. To express variant subtilisins, plasmids were transformed into a subtilisin-neutral protease deletion mutant of B. subtilis, BG2036, as previously described. Cultures were fermented in shake flasks for 24 h at 37°C in LB media containing 12.5 mg/mL chloramphenicol and subtilisin was purified from culture supernatants as described. Purity of subtilisin was greater than 95% as judged by SDS PAGE.

These mutant plasmids designated pS156 and pQ156 and mutant subtilisins designated S156 and Q156 were analyzed with the above synthetic substrates where P-1 comprised the amino acids Glu, Gln, Met and Lys. The results of this analyses are presented in Example 9.

EXAMPLE 9

Multiple Mutants With Altered Substrate Specificity - Substitution at Positions 156 and 166

Single substitutions of position 166 are described in Examples 3 and 4. Example 8 describes single substitutions at position 156 as well as the protocol of Figure 21 whereby various double mutants comprising the substitution of various amino acids at positions 156 and 166 can be made. This example describes the construction and substrate specificity of subtilisin containing substitutions at position 156 and 166 and summarizes some of the data for single and double mutants at positions 156 and 166 with various substrates.

K166 is a common replacement amino acid in the 156/166 mutants described herein. The replacement of Lys for

Gly166 was achieved by using the synthetic DNA cassette at the top right of Figure 21 which contained the triplet AAA for NNN. This produced fragment 2 with Lys substituting for Gly166.

The 156 substituents were Gln and Ser. The Gln and Ser substitutions at Gly156 are contained within fragment 3 (bottom right Figure 21).

The multiple mutants were produced by combining fragments 1, 2 and 3 as described in Example 8. The mutants Q156/K166 and S156/K166 were selectively generated by differential phosphorylation as described. Alternatively, the double 156/166 mutants, c.f. Q156/K166 and S156/K166, were prepared by ligation of the 4.6kb SacI-BamHI fragment from the relevant p156 plasmid containing the 0.6kb SacI-BamHI fragment from the relevant p166 plasmid.

These mutants, the single mutant K166, and the S156 and Q156 mutants of Example 8 were analyzed for substitute specificity against synthetic polypeptides containing Phe or Glu as the P-1 substrate residue. The results are presented in Table XIII.

TABLE XIII

Enzymes Compared[b]	Substrate P-1 Residue	kcat	Km	kcat/Km	kcat/Km (mutant) / kcat/Km(wt)
Glu156/Gly166 (WT)	Phe	50.00	1.4×10^{-4}	3.6×10^{5}	(1)
	Glu	0.54	3.4×10^{-2}	1.6×10^{1}	(1)
K166	Phe	20.00	4.0×10^{-5}	5.2×10^{5}	1.4
	Glu	0.70	5.6×10^{-5}	1.2×10^{4}	750
Q156/K166	Phe	30.00	1.9×10^{-5}	1.6×10^{6}	4.4
	Glu	1.60	3.1×10^{-5}	5.0×10^{4}	3100
S156/K166	Phe	30.00	1.8×10^{-5}	1.6×10^{6}	4.4
	Glu	0.60	3.9×10^{-5}	1.6×10^{4}	1000
S156	Phe	34.00	4.7×10^{-5}	7.3×10^{5}	2.0
	Glu	0.40	1.8×10^{-3}	1.1×10^{2}	6.9
E156	Phe	48.00	4.5×10^{-5}	1.1×10^{6}	3.1
	Glu	0.90	3.3×10^{-3}	2.7×10^{2}	17

As can be seen in Table XIV, either of these single mutations improve enzyme performance upon substrates with glutamate at the P-1 enzyme binding site. When these single mutations were combined, the resulting multiple enzyme mutants are better than either parent. These single or multiple mutations also alter the relative pH activity profiles of the enzymes as shown in Figure 23.

To isolate the contribution of electrostatics to substrate specificity from other chemical binding forces, these various single and double mutants were analyzed for their ability to bind and cleave synthetic substrates containing Glu, Gln, Met and Lys as the P-1 substrate amino acid. This permitted comparisons between side-chains that were more sterically similar but differed in charge (e.g., Glu versus Gln, Lys versus Met). Similarly, mutant enzymes were assayed against homologous P-1 substrates that were most sterically similar but differed in charge (Table XIV).

TABLE XIV

Kinetics of Position 156/166 Subtilisins
Determined for Different P1 Substrates

Enzyme Position [a]		Net Charge [b]	P-1 Substrate log kcat/Km (log 1/Km) [c]							
			Glu		Gln		Met		Lys	
156	166									
Glu	Asp	-2	n.d.		3.02	(2.56)	3.93	(2.74)	4.23	(3.00)
Glu	Glu	-2	n.d.		3.06	(2.91)	3.86	(3.28)	4.48	(3.69)
Glu	Asn	-1	1.62	(2.22)	3.85	(3.14)	4.99	(3.85)	4.15	(2.88)
Glu	Gln	-1	1.20	(2.12)	4.36	(3.64)	5.43	(4.36)	4.10	(3.15)
Gln	Asp	-1	1.30	(1.79)	3.40	(3.08)	4.94	(3.87)	4.41	(3.22)
Ser	Asp	-1	1.23	(2.13)	3.41	(3.09)	4.67	(3.68)	4.24	(3.07)
Glu	Met	-1	1.20	(2.30)	3.89	(3.19)	5.64	(4.83)	4.70	(3.89)
Glu	Ala	-1	n.d.		4.34	(3.55)	5.65	(4.46)	4.90	(3.24)
Glu	Gly(wt)	-1	1.20	(1.47)	3.85	(3.35)	5.07	(3.97)	4.60	(3.13)
Gln	Gly	0	2.42	(2.48)	4.53	(3.81)	5.77	(4.61)	3.76	(2.82)
Ser	Gly	0	2.31	(2.73)	4.09	(3.68)	5.61	(4.55)	3.46	(2.74)
Gln	Asn	0	2.04	(2.72)	4.51	(3.76)	5.79	(4.66)	3.75	(2.74)
Ser	Asn	0	1.91	(2.78)	4.57	(3.82)	5.72	(4.64)	3.68	(2.80)
Glu	Arg	0	2.91	(3.30)	4.26	(3.50)	5.32	(4.22)	3.19	(2.80)
Glu	Lys	0	4.09	(4.25)	4.70	(3.88)	6.15	(4.45)	4.23	(2.93)
Gln	Lys	+1	4.70	(4.50)	4.64	(3.68)	5.97	(4.68)	3.23	(2.75)
Ser	Lys	+1	4.21	(4.40)	4.84	(3.94)	6.16	(4.90)	3.73	(2.84)
Maximum difference:										
log kcat/Km (log 1/Km) [d]			3.5	(3.0)	1.8	(1.4)	2.3	(2.2)	-1.3	(-1.0)

Footnotes to Table XIV:

(a) B. subtilis, BG 2036, expressing indicated variant subtilisin were fermented and enzymes purified as previously described (Estell, et al. (1985) J. Biol. Chem. 260, 6518-6521). Wild type subtilisin is indicated (wt) containing Glu156 and Gly166.

(b) Net charge in the P-1 binding site is defined as the sum of charges from positions 156 and 166 at pH 8.6.

(c) Values for kcat(s^{-1}) and Km(M) were measured in 0.1M Tris pH 8.6 at 25°C as previously described against P-1 substrates having the form succinyl-L-AlaL-AlaL-ProL-[X]-p-nitroanilide, where X is the indicated P-1 amino acid. Values for log 1/Km are shown inside parentheses. All errors in determination of kcat/Km and 1/Km are below 5%.

(d) Because values for Glu156/Asp166(D166) are too small to determine accurately, the maximum difference taken for GluP-1 substrate is limited to a charge range of +1 to -1 charge change.

n.d. = not determined

The kcat/Km ratios shown are the second order rate constants for the conversion of substrate to product, and represent the catalytic efficiency of the enzyme. These ratios are presented in logarithmic form to scale the data, and because log kcat/Km is proportional to the lowering of transition-state activation energy (ΔG_T). Mutations at position 156 and 166 produce changes in catalytic efficiency toward Glu, Gln, Met and Lys P-1 substrates of 3100, 60, 200 and 20 fold, respectively. Making the P-1 binding-site more positively charged [e.g., compare Gln156/Lys166 (Q156/K166) versus Glu156/Met166 (Glu156/M166)] dramatically increased kcat/Km toward the Glu P-1 substrate (up to 3100 fold), and decreased the catalytic efficiency toward the Lys P-1 substrate (up to 10 fold). In addition, the results show that the catalytic efficiency of wild type enzyme can be

greatly improved toward any of the four P-1 substrates by mutagenesis of the P-1 binding site.

The changes in kcat/Km are caused predominantly by changes in 1/Km. Because 1/Km is approximately equal to 1/Ks, the enzyme-substrate association constant, the mutations primarily cause a change in substrate binding. These mutations produce smaller effects on kcat that run parallel to the effects on 1/Km. The changes in kcat suggest either an alteration in binding in the P-1 binding site in going from the Michaelis-complex E·S) to the transition-state complex (E-S≠) as previously proposed (Robertus, J.D., et al. (1972) Biochemistry 11, 2439-2449; Robertus, J.D., et al. (1972) Biochemistry 11, 4293-4303), or change in the position of the scissile peptide bond over the catalytic serine in the E*S complex.

Changes in substrate preference that arise from changes in the net charge in the P-1 binding site show trends that are best accounted for by electrostatic effects (Figure 28). As the P-1 binding cleft becomes more positively charged, the average catalytic efficiency increases much more for the Glu P-1 substrate than for its neutral and isosteric P-1 homolog, Gln (Figure 28A). Furthermore, at the positive extreme both substrates have nearly identical catalytic efficiencies.

In contrast, as the P-1 site becomes more positively charged the catalytic efficiency toward the Lys P-1 substrate decreases, and diverges sharply from its neutral and isosteric homolog, Met (Figure 28B). The similar and parallel upward trend seen with increasing positive charge for the Met and Glu P-1 substrates probably results from the fact that all the substrates

are succinylated on their amino-terminal end, and thus carry a formal negative charge.

The trends observed in log kcat/Km are dominated by changes in the Km term (Figures 28C and 28D). As the pocket becomes more positively charged, the log 1/Km values converge for Glu and Gln P-1 substrates (Figure 28C), and diverge for Lys and Met P-1 substrates (Figure 28D). Although less pronounced effects are seen in log kcat, the effects of P-1 charge on log kcat parallel those seen in log 1/Km and become larger as the P-1 pocket becomes more positively charged. This may result from the fact that the transition-state is a tetrahedral anion, and a net positive charge in the enzyme may serve to provide some added stabilization to the transition-state.

The effect of the change in P-1 binding-site charge on substrate preference can be estimated from the differences in slopes between the charged and neutral isosteric P-1 substrates (Figure 28B). The average change in substrate preference (Δlog kcat/Km) between charged and neutral isosteric substrates increases roughly 10-fold as the complementary charge or the enzyme increases (Table XV). When comparing Glu versus Lys, this difference is 100-fold and the change in substrate preference appears predominantly in the Km term.

TABLE XV

Differential Effect on Binding Site Charge on log kcat/Km or (log 1/Km) for P-1 Substrates that Differ in Charge[a]

Change in P-1 Binding Site Charge[b]	Δlog kcat/Km GluGln	Δlog kcat/Km MetLys	(Δlog 1/Km) GluLys
-2 to -1	n.d.	1.2 (1.2)	n.d.
-1 to 0	0.7 (0.6)	1.3 (0.8)	2.1 (1.4)
0 to +1	1.5 (1.3)	0.5 (0.3)	2.0 (1.5)
Avg. change in log kcat/K_m or (log 1/Km) per unit charge change	1.1 (1.0)	1.0 (0.8)	2.1 (1.5)

(a) The difference in the slopes of curves were taken between the P-1 substrates over the charge interval given for log (kcat/Km) (Figure 28A, B) and (log 1/Km) (Figure 28C, D). Values represent the differential effect a charge change has in distinguishing the substrates that are compared.

(b) Charge in P-1 binding site is defined as the sum of charges from positions 156 and 166.

The free energy of electrostatic interactions in the structure and energetics of salt-bridge formation depends on the distance between the charges and the microscopic dielectric of the media. To dissect these structural and microenvironmental effects, the energies involved in specific salt-bridges were evaluated. In addition to the possible salt-bridges shown (Figures 29A and 29B), reasonable salt-bridges can be built between a Lys P-1 substrate and Asp at position 166, and between a Glu P-1 substrate and a Lys at position 166 (not shown). Although only one of these structures is confirmed by X-ray crystalography (Poulos, T.L., et al. (1976) J. Mol. Biol. 257 1097-1103), all models have favorable torsion angles (Sielecki, A.R., et al. (1979) J. Mol. Biol. 134, 781-804), and do not introduce unfavorable van der Waals contacts.

The change in charged P-1 substrate preference brought about by formation of the model salt-bridges above are shown in Table XVI.

TABLE XVI

Effect of Salt Bridge Formation Between Enzyme and Substrate on P1 Substrate Preference[a]

Enzymes Compared[b] 1	Enzymes Compared[b] 2	Enzyme Position Changed	P-1 Substrates Compared	Substrate[d] Preference Δlog (kcat/Km) 1	Substrate[d] Preference Δlog (kcat/Km) 2	Change in Substrate Preference ΔΔlog (kcat/Km) (1-2)
Glu156/Asp166	Gln156/Asp166	156	LysMet	+0.30	-0.53	0.83
Glu156/Asn166	Gln156/Asn166	156	LysMet	-0.84	-2.04	1.20
Glu156/Gly166	Gln156/Gly166	156	LysMet	-0.47	-2.10	1.63
Glu156/Ley-166	Gln156/Lys166	156	LysMet	-1.92	-2.74	0.82
				Ave ΔΔlog (kcat/Km)		1.10 ± 0.3
Glu156/Asp166	Glu156/Asn166	166	LysMet	+0.30	-0.84	1.14
Glu156/Glu166	Glu156/Glu166	166	LysMet	+0.62	-1.33	1.95
Gln156/Asp166	Gln156/Asn166	166	LysMet	-0.53	-2.04	1.51
Ser156/Asp166	Ser156/Asn166	166	LysMet	-0.43	-2.04	1.61
Glu156/Lys166	Glu156/Met166	166	GluGln	-0.63	-2.69	2/06
				Ave ΔΔlog (kcat/Km)		1.70 ± 0.3

Footnotes to Table XVI:

(a) Molecular modeling shows it is possible to form a salt bridge between the indicated charged P-1 substrate and a complementary charge in the P-1 binding site of the enzyme at the indicated position changed.

(b) Enzymes compared have sterically similar amino acid substitutions that differ in charge at the indicated position.

(c) The P-1 substrates compared are structurally similar but differ in charge. The charged P-1 substrate is complementary to the charge change at the position indicated between enzymes 1 and 2.

(d) Date from Table XIV was used to compute the difference in log (kcat/Km) between the charged and the non-charged P-1 substrate (i.e., the substrate preference). The substrate preference is shown separately for enzyme 1 and 2.

(e) The difference in substrate preference between enzyme 1 (more highly charged) and enzyme 2 (more neutral) represents the rate change accompanying the electrostatic interaction.

The difference between catalytic efficiencies (i.e., Δlog kcat/Km) for the charged and neutral P-1 substrates (e.g., Lys minus Met or Glu minus Gln) give the substrate preference for each enzyme. The change in substrate preference (ΔΔlog kcat/Km) between the charged and more neutral enzyme homologs (e.g., Glu156/Gly166 minus Gln156(Q156)/Gly166) reflects the change in catalytic efficiency that may be attributed solely to electrostatic effects.

These results show that the average change in substrate preference is considerably greater when electrostatic substitutions are produced at position 166 (50-fold in kcat/Km) versus position 156 (12-fold in kcat/Km). From these ΔΔlog kcat/Km values, an average change in transition-state stabilization energy can be calculated of -1.5 and -2.4 kcal/mol for

substitutions at positions 156 and 166, respectively. This should represent the stabilization energy contributed from a favorable electrostatic interaction for the binding of free enzyme and substrate to form the transition-state complex.

EXAMPLE 10

Substitutions at Position 217

Tyr217 has been substituted by all other 19 amino acids. Cassette mutagenesis as described in EPO publication No. 0130756 was used according to the protocol of Figure 22. The EcoRV restriction site was used for restriction-purification of pA217.

Since this position is involved in substrate binding, mutations here effect kinetic parameters of the enzyme. An example is the substitution of Leu for Tyr at position 217. For the substrate sAAPFpNa, this mutant has a kcat of 277 5' and a Km of 4.7×10^{-4} with a kcat/Km ratio of 6×10^{5}. This represents a 5.5-fold increase in kcat with a 3-fold increase in Km over the wild type enzyme.

In addition, replacement of Tyr217 by Lys, Arg, Phe or Leu results in mutant enzymes which are more stable at pHs of about 9-11 than the WT enzyme. Conversely, replacement of Tyr217 by Asp, Glu, Gly or Pro results in enzymes which are less stable at pHs of about 9-11 than the WT enzyme.

EXAMPLE 11

Multiple Mutants Having Altered Thermal Stability

B. amyloliquefacien subtilisin does not contain any cysteine residues. Thus, any attempt to produce thermal stability by Cys cross-linkage required the substitution of more than one amino acid in subtilisin with Cys. The following subtilisin residues were multiply substituted with cysteine:

Thr22/Ser87

Ser24/Ser87

Mutagenesis of Ser24 to Cys was carried out with a 5' phosphorylated oligonucleotide primer having the sequence

5'-pC-TAC-ACT-GGA-TG**C-AAT-GTT-AAA-G-3',

(Asterisks show the location of mismatches and the underlined sequence shows the position of the altered Sau3A site.) The B. amyloliquefaciens subtilisin gene on a 1.5 kb EcoRI-BAMHI fragment from pS4.5 was cloned into M13mp11 and single stranded DNA was isolated. This template (M13mp11SUBT) was double primed with the 5' phosphorylated M13 universal sequencing primer and the mutagenesis primer. Adelman, et al. (1983) DNA 2, 183-193. The heteroduplex was transfected into competent JM101 cells and plaques were probed for the mutant sequence (Zoller, M.J., et al. (1982) Nucleic Acid Res. 10, 6487-6500; Wallace, et al. (1981) Nucleic Acid Res. 9, 3647-3656) using a tetramethylammonium chloride hybridization protocol (Wood, et al. (1985) Proc. Natl. Acad. Sci. USA 82, 1585-1588). The Ser87 to Cys mutation was prepared in